ARI Research Note 97-34

Effects of Stress on Judgment and Decision Making

Kenneth R. Hammond and James K. Doyle

Research and Advanced Concepts Office
Michael Drillings, Chief

September 1997

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Battelle

Technical review by

Michael Drillings

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED (from. . . to)
1997, September	Final	May 1991-November 1991

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Effects of Stress on Judgment and Decision Making	DAAL03-86-D-0001 DO 2715
	5b. PROGRAM ELEMENT NUMBER 0601102A
6. AUTHOR(S)	5c. PROJECT NUMBER B74F
Kenneth R. Hammond and James K. Doyle	5d. TASK NUMBER 1901
	5e. WORK UNIT NUMBER C51

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Dr. Kenneth R. Hammond 1740 Columbine Avenue Boulder, CO 80302	TCN 91443
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-BR 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Note 97-34

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Task was performed under a Scientific Services Agreement issued by Battelle, Research Triangle Park Office, 200 Park Drive, P.O. Box 12297, Research Triangle Park, NC.

14. ABSTRACT *(Maximum 200 words)*:

This monograph (Part II) is the second of a planned three-part series. Following Part I, which examines four literatures related to judgment and decision making (J/DM) under stress, Part II narrows its focus to a detailed treatment of stress within the J/DM literature. Six sections are included: (a) an introduction, bridging Parts I and II, (b) a consideration of two principal topics (rationality and performance) in the J/DM literature in relation to stress, (c) an examination of current textbooks, anthologies, and reviews, as well as books and articles in the human factors field, with regard to their treatment of stress and J/DM, (d) a description of current J/DM theories and models and their potential utility for the study of J/DM under stress, (e) an examination of methodological issues bearing on research on J/DM under stress, and (f) the outline of a new approach intended to advance theory and method.

The general conclusion drawn from the examination of the aforementioned material (including an additional several hundred articles not cited) is the same as that drawn from the work in Part I, namely, research is lacking in coherent theoretical background, diffuse in content, and completely lacking in secure generalizations.

If progress is to be made with regard to this topic, critically important to the military and other sections of society, a resolute, comprehensive effort will have to be made, theoretically, methodologically, and empirically. Part III (forthcoming) will consist of an attempt to meet these goals.

15. SUBJECT TERMS

Stress Judgment Decision making

Rationality Performance

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	163	

Preface

This monograph constitutes Part II of a three-part series. Its primary aim is to bring the studies of the effects of stress on judgment and decision making into sharper focus than was possible in Part I, which reviewed *four* literatures.

The monograph is organized as follows: We first mention the conclusions drawn from the study of four literatures in Part I, then consider the two main topics that interest researchers in the field of J/DM, examine the manner in which several textbooks in the J/DM field and the human factors field treat the topic of stress effects on J/DM, provide brief descriptions of several stable J/DM research programs, consider certain methodological issues, draw overall conclusions about what we do and do not know about the effects of stress on J/DM, and finally, provide a brief outline of a new approach to this topic.

The reader will note that we often provide lengthy quotations from various authors. That is because the monograph is essentially a review of the work of others. In the interests of allowing the reader to grasp the arguments put forward in the quotations we have deliberately erred on the side of including more, perhaps, than usual. We are confident, however, that the authors of the quotations will admire our judgment.

Not every article that we acquired and read has been included in the text either in a substantive form or even as a citation; there simply was not time to do so. We have, however, provided a bibliography that

includes most of the articles we examined and considered relevant to the topic of stress and J/DM.

We thank Mary Luhring and Doreen Petersen for their assistance—without which we would not have even considered undertaking this endeavor.

This work was supported by the Army Research Institute for Behavioral and Social Sciences (Dr. George W. Lawton) under the auspices of the U.S. Army Research Office Scientific Services Program administered by Battelle (Delivery Order 2715, Contract No. DAAL03-86-D-0001).

Table of Contents

EFFECTS OF STRESS ON JUDGMENT AND DECISION MAKING

1. Introduction

The foregoing review (see Part I) indicates that, aside from the voluminous growth of the literature on stress, formidable problems face researchers in this field; they range from the most elementary—clarifying and defining the concept of stress—to the most complex—reconciling the content of four largely independent literatures, and reducing the isolation of theory and research between and within those literatures. Corroboration of this conclusion can be found in Zajonc and Markus (1984), who also noticed and remarked on this isolation:

> In contemporary psychology, cognitive and affective processes are treated within largely separate and distinct conceptual frameworks, and, with few exceptions (e.g., Lang, 1979; Mandler, 1975), scientific publications in one area of research do not cite those in the other. Yet both domains of research investigate processes that interact with one another constantly and vigorously. Even though most theories of emotion assume as necessary the extensive participation of cognitive functions (Lazarus, 1966; Mandler, 1975; Schachter & Singer, 1962), the precise nature of this participation has been seldom explicitly analyzed. And it is equally remarkable that, even though cognitive content is rarely processed without the participation of affect (Piaget, 1981), cognitive theories have no conceptual elements that reflect the contribution of affective factors (Zajonc, 1980). This conceptual isolation of affect and cognition is likely to persist unless we come to understand which elements of these two processes make contact with each other and how the influence of one process over the other is actually effected. (p. 73)

There is, however, no evidence to suggest that the content of the *four* literatures described in Part I will ever be reconciled—with the possible exception of the physiological and ergonomic literatures. (For an attempt to integrate the latter two, see Hockey, Gaillard, & Coles, 1986.) Many anthologies purport to bring together studies in

this field (e.g., Spielberger and Sarason alone have published 12 anthologies from 1975 to 1989 on "Stress and Anxiety"), but persistent, systematic, and cumulative research has not become apparent. Even the military has shown only sporadic and unsustained interest in research on this topic despite its obvious significance for all phases of military activity, and command and control in particular. So far as we know, there is no compendium of information that is used, or can be used, for anticipating and coping with the effects of stress on judgment.

In short, our knowledge remains fragmentary; the clinical/social psychological literature, the physiological literature, and the human factors/ergonomics literatures do not yet tell us whether specific environmental (exogenous) conditions, generally assumed to be stressors (time pressure, sleep loss, heat, noise, etc.), have specific or general effects on rationality or performance, at what point their effects can be expected, or in any case what these effects will be. Nor do we yet have clear evidence that stress, once induced and independently confirmed, has any regularly dependable effects on the topics or processes involved in judgment and decision making.

Research Strategy: Time for Change

It is now appropriate to change the research strategy from the current one if for no other reason than that it is not working. Thousands of articles have been published on the topic of stress and cognitive functioning for decades, yet its most experienced researchers report that firm conclusions remain elusive. Current research strategy—seeking first to ascertain the effects of various

stressors (e.g., temperature, heat) on a variety of psychological functions (e.g., attention) and then attempting, somehow, to cumulate the results into coherent conclusions is certainly a plausible research activity—simply follows the general pattern of research in psychology. But neither cumulative results nor theoretical coherence has been achieved, either *because* of the pervasiveness of conventional research methodology (our view) or *in spite of it* (as most psychologists would believe). In our view, the conventional research strategy has had its chance; it is time for a new approach.

We are not alone in this conclusion. The human factors researchers Hancock and Warm (1989) in commenting on research on vigilance assert:

> There has been a collective failure of theories that seek to explain vigilance performance (see Loeb and Alluisi, 1984). This failure is also true for theories of stress in general, which with few exceptions have exhibited similar stagnation. It is noteworthy that the only theoretical construct that spans the two areas is the concept of behavioral arousal. In their paper Koelega et al. (1986) observed, "But arousal theory can explain any results, post hoc, and lacks predictive power. The position on the inverted-U curve can only be specified after the experiment, so arousal theory, in its present form, is not amenable to rigorous experimental testing" (p. 588). They are assuredly correct. This and additional limitations of the unitary behavioral arousal theory have been elaborated in detail by Hancock (1987). Failure to find consistencies in the noise and vigilance data is consequently a specific case of the general failure of theoretical integration both within and across two respective areas. . . .
>
> At present no satisfactory theoretical account is available to predict the action of discrete or interactive stresses that occur in real-world settings (see Hockey et al., 1986). (pp. 524-525)

Moray, also an experienced researcher in the human factors field, expresses strong reservations about current methodology:

> Human error probability (HEP) . . . shows very great fluctuations due to such factors as circadian rhythms, health, stress, etc., and . . . it will fluctuate in addition as a function of the dynamics of task demands, since attention, and hence information acquisition, is tightly coupled to the dynamics of the task. How best to incorporate human error models into PRA [probabilistic risk assessment] is an unsolved problem, but *certainly it cannot be done in the way which has been attempted so far* [italics added], . . . which tries to provide a fixed point estimate with an order of magnitude range for a tactically variable, task-coupled, adaptive class of meaningful behaviour. (1990, pp. 1211-1212)

(See also Simon (1979, pp. 320-327) on "testing theories statistically" for negative views on conventional methodology.) In sum, strong calls for methodological change are appearing among those most experienced in the empirical analysis of the effects of stress on cognitive functioning.

The Field of Judgment and Decision Making

If these calls for change are accepted—and we accept them—then one tactic to pursue would be to turn to Literature IV—the field of Judgment and Decision Making (J/DM)—for theoretical orientation, methodological guidance, and empirical results relevant to understanding judgment under stress. But this tactic might well be met with disappointment; there is as yet no unified theory of J/DM (but see attempts by Hammond, 1980, 1986b), this literature remains largely dependent upon the same conventional methodology that dominates psychology, and there are few firm empirical results that bear on the topic of stress and judgment. Nevertheless, there are clearly identifiable, systematic approaches in the J/DM field that have persisted and have continued to be cumulatively productive for 15-30 years, despite being hampered by vestiges of conventional methodology (such as between-subject analyses). Each one of these

approaches can be described as a "progressive scientific research program" (to use Lakatos's, 1981, term) that is at least partially successful. For example, work by N. Anderson (1981), Einhorn and Hogarth (particularly on judging probable cause, 1986), Hammond, Brehmer, and others within the Brunswikian tradition (Brehmer & Joyce, 1988; Brunswik, 1956; Hammond, 1988), Kahneman, Slovic, and Tversky (1982 et seq.), Payne and his colleagues (Payne, Bettman, & Johnson, 1992), and von Winterfeldt and Edwards (1986) as well as that of Simon, and recently that of John Anderson (1990), is sufficiently systematic and cumulative that each offers a definite point of departure for examining the effects of stress. Although there are persistent differences among these approaches, their focus on certain common topics is sufficient to bring these researchers and many others together in annual meetings and common journal publications. Significantly, several textbooks as well as books of readings have appeared.

The research programs mentioned above thus offer an opportunity for the research strategy Hockey (1979) correctly called for, namely, a focus on descriptions of how stress induces—if it does—"fundamental changes in [cognitive] functioning." These research programs now provide the the best source of parameters of cognitive processes, and the best source of hypotheses about "fundamental changes" in J/DM likely to be induced by stress. Evidence for this argument can be seen in the fact that for the first time textbooks on J/DM resemble one another; they include roughly the same material. Although they have little to say about stress, they discuss many of the

same topics, describe and criticize the same theories, and describe the various methods used in various programs. (See, for example, textbooks by Baron, 1988; Bazerman, 1990; Dawes, 1988; Hogarth, 1987; Rachlin, 1989; Russo & Schoemaker, 1989; von Winterfeldt & Edwards, 1986; Watson & Buede; 1987, Yates, 1990; see also the review of textbooks by Hastie, 1991; see also anthologies by Arkes & Hammond, 1986; Kahneman et al., 1982; Rohrmann, Beach, Vlek, & Watson, 1989; and recent reviews by Brehmer & Joyce, 1988; Payne et al., 1992.) Although the clear differences in the conceptual frameworks that guide these research programs remain much the same as when described by Hammond, McClelland, and Mumpower in 1980, and thus suggest impermeable barriers, on the other hand, the persistence of differences offers evidence of the definitive nature of each program. (Of course, resolution of differences might well lead to greater advancement.)

Perhaps the most important and most useful product of these research programs is a distinction drawn by all of them, namely, the distinction between *prescriptions for* and *descriptions of* cognitive activity. Offering *prescriptions*, that is, specific standards for judgments and decisions, constitutes a unique and positive contribution from these programs. And offering *descriptions* means offering models of the processes that account for the judgments observed. Work outside the field of J/DM rarely offers either.

Prescriptive Standards for Evaluating Stress Effects

Prescriptions for judgment and decision making are offered in the form of standards for rationality or performance.

Normative decision rules can provide a criterion, a standard, by which rationality is evaluated under benign or stressful conditions. Thus, for example, Bayes' Rule can provide a baseline for ascertaining the effects of stress, surely a great convenience. Perhaps most important, current textbooks reflect the conclusions drawn by many researchers that few persons achieve normative rules of rationality, even under the most benign circumstances—even when encouraged to do so. If it is true that such standards are rarely achieved under benign conditions, then some new criterion for rationality will have to be found, more easily said than done. But even if normative criteria are seldom met (still a matter of debate), prescriptions for cognitive activity are bound to have some use in the measurement of stress effects inasmuch as they offer a *form* of cognition against which cognitive activity can be evaluated. That is, stress may result in a *decreased* deformation of a prescriptive rule, or it may result in certain predictable *increases* in deformation. Perhaps because most empirical research on the effects of stress on cognitive activity has been carried out by human factors researchers with a strong interest in performance, no studies, so far as we can ascertain, have investigated the effects of stress on rationality, however described.

In the case of *performance*, empirical accuracy, however cognitively achieved, prescribes the ultimate baseline; departures therefrom can be evaluated in terms of absolutes when errors in accuracy are intolerable (e.g., in the case of operators of nuclear plants) or in terms of reasonable standards when errors are tolerated (e.g., stock brokers, weather forecasters). Because of their

concreteness, empirical measures of performance are attractive criteria for evaluating stress effects that do not suffer the disadvantages of disputes about *which* standard of rationality should apply, or how any standard *should* apply. Thus, it is not surprising that empirical accuracy has been almost universally employed—particularly by human factors researchers—as a standard against which to measure stress effects.

Some of the current research programs have focused their efforts on observations of departures from rationality (to some extent the research program associated with Edwards and von Winterfeldt [von Winterfeldt & Edwards, 1986] and, to a large extent, the research program associated with Tversky and Kahneman [Kahneman et al., 1982]), each with a different point of view. Others have focused on performance (e.g., Brehmer and Hammond [Brehmer & Joyce, 1988]). Thus it is not surprising to find that various research programs have developed different *descriptive* models. Differences in these models, and their associated methodology, are so large we devote one section of this report to them (see below).

In sum, although the field of J/DM does not offer the stress researcher a homogeneous field of endeavor, unified in theory and methodology, it does offer a set of strong research programs, each persistent and cumulative over the past two or three decades, and each of them offers (a) theoretical models of the judgment process that will provide parameters potentially susceptible to stressors, and (b) a source of hypotheses regarding the likely effects of such stressors.

Current Assumptions Regarding the Effects of Stress on J/DM

It is now generally assumed that stressful conditions will almost always have a disruptive, negative effect on judgment and decision making (there is little evidence to the contrary). Two specific effects are to be expected: (a) departures from standards of rationality (incoherence), and (b) degraded performance (errors, or lack of correspondence with a criterion). The first effect (incoherence) is generally assumed to be due to the loss of trained, analytical, rational, perform-to-rule cognition (rather than, say, Bayesian incoherence) and the subsequent turn to error-filled intuitive cognition produced by heuristics, cognitive biases, logical fallacies, computational errors, perceptual errors, and even stupidity. But these putative effects of stress have seldom, if ever, been directly addressed within the context of normative models; we know virtually nothing about such matters. The second empirical effect (empirical error) is assumed to result from any of the above as well as a number of physiological causes (e.g., arousal is the favorite, despite the growing doubts indicated by Hancock & Warm, 1989, and others) and psychological causes (see Literatures I, II, and III). Although these causes have frequently been studied with respect to performance, they have yet to be systematically examined in the context of the modern J/DM research programs mentioned above.

Summary

The research programs in the field of J/DM encompass the principal contemporary efforts toward understanding judgment and decision processes. Thus, they offer the most advantageous point of

departure for investigating the effects of stress on both rationality and performance. Therefore, in what follows we consider the *topics* addressed by various research programs, show how the topic of stress has been treated in recent *textbooks and articles*, briefly describe current *research programs* and indicate how the topic of stress might be treated within them, and examine and criticize current *methodologies* (usually implicit) that support the research and methods (usually explicit) employed.

In short, Part II of this monograph brings the main topics, theories, and methods that have been developed within the major *current* research programs in Literature IV—the literature of judgment and decision making—to bear on the topic of stress. Part III (forthcoming) will present a new approach to this topic.

2. Current Topics in J/DM Related to Stress Research

Judgment and Decision Making

This section is devoted to describing the two principal topics in the J/DM field: (a) cognitive *coherence*, or rationality, of a person's judgments and decisions, and (b) cognitive *performance*, the empirical accuracy (or correspondence with an empirical criterion) of a person's judgments and decisions.

Laudan's (1981) discussion of this distinction shows that it is not peculiar to the J/DM field, rather it is a reflection of a broader "tension" in the history of science:

> Running through much of the history of the philosophy of science is a tension between coherentist and correspondentist accounts of scientific knowledge. Coherentists stress the need for appropriate types of conceptual linkages between our beliefs, while correspondentists emphasize the grounding of beliefs in the world. Each account typically makes only minimal concessions to the other. (Correspondentists, for instance, will usually grant that theories should minimally cohere in the sense of being consistent with our other beliefs.) Neither side, however, has been willing to grant that a *broad range* of both empirical and conceptual checks are of equal importance in theory testing. (p. 146)

Researchers in the field of J/DM can also be classified as pursuing these two different "accounts" of J/DM and have also maintained the distance from one another that Laudan has noted in the history of science. Moreover, they do not hesitate to generalize their results from one type of research to the other—a practice noted by Slovic and Lichtenstein in 1971, and one that continues today.

Vicente (1990) has stressed the practical importance of the correspondence/coherence distinction in connection with system design:

> A distinction is made between coherence- and correspondence-driven work domains. This novel domain taxonomy is used to argue that the widely accepted goal of making the interface representation compatible with the user's mental model is not always appropriate. For correspondence-driven domains, it is more meaningful to constrain design from the side of the work domain rather than from that of the user. The implications of the coherence/correspondence distinction for the modelling of work domains, for interface design in computer supported co-operative work, and for the development of a multidimensional taxonomy of work domains are also briefly pointed out. The discussion suggests that the correspondence/coherence taxonomy provides a powerful conceptual tool for addressing fundamental issues in human-computer interaction. (p. 493)

Vicente is a design engineer and therefore interested in which type of interface between the environment and operator is most likely to be supportive when the unexpected occurs, that is, when stress is induced. His recognition of the distinction between coherence and correspondence is therefore important for the future of stress research in the workplace. Thus the broad "tension" Laudan (1981) finds in the history of the philosophy of science maintains its place—under the most practical of applications.

Are Human Beings Optimal or Poor Decision Makers?

Schoemaker (1991) has recently given prominence to this question by his target article in *The Behavioral and Brain Sciences* entitled "The Quest for Optimality:" A Positive Heuristic of Science?" Schoemaker (p. 205) observes that "the optimality approach is prone to systematic biases," thus lending even greater force to the ambiguity of the nature of the concept of optimality and, in turn, its doubtful place in the evaluation of the effect of stress.

John Anderson's (1991b) comments on Schoemaker's excellent and useful article are informative because they illustrate the difference

between optimal performance from a rational/competence ("coherentist") point of view and of an empirical performance (correspondence) point of view:

> To explore the possibility of an optimality explanation of memory, we needed a theory of the information-processing demands placed on memory by the environment. Adapting theories developed for library borrowing (Burrell & Cane 1982) and file access (Stritter 1977) we (Anderson & Milson 1989) showed that many memory phenomena could be seen as optimal responses to the statistical structure of retrieval requests from the environment. More specifically, we showed that human memory displays the fastest retrieval latencies and highest probability of recall for the information that is statistically most likely to be needed.
>
> This research was based on information-retrieval demands placed on nonhuman systems, however. More recently, Lael Schooler and I set out to study carefully the actual information processing demands placed on humans. We looked at a number of such computerized sources of input to humans as topics in the *New York Times*, electronic mail messages, and words spoken to young children. This is not the place to describe our results in detail, but in every case we found that memory functions mirrored perfectly the statistical properties in the environment. (p. 215)

(Readers should note the parallel conclusions drawn from many studies of multiple cue probability learning, although Anderson does not cite them; see, for example, Brehmer & Joyce, 1988; and Hammond et al., 1980.) Anderson continues:

> Thus, we are finding, contrary to all expectations, that human memory seems exquisitely tuned to the statistics of information presentation in the environment. Whether this will ultimately be viewed as a case of optimization or it will be given some other explanation, it illustrates the potential for optimality considerations to lead to novel insight. This would never have been known unless we chose to penetrate beyond the apparent nonoptimality of human memory. (Anderson, 1991b, p. 215)

In short, Anderson comes down unequivocally on the "optimal" and "correspondence" side of the division.

Anderson's views are presented in detail in his book (1990), the title of which *(The Adaptive Character of Thought)* because of Anderson's stature, illustrates a significant change in cognitive psychology in the direction of the correspondence view. Anderson's views are also presented in an article in *The Behavioral and Brain Sciences* (1991a). This article and the many commentaries that follow, as well as Anderson's reply to them, carry considerable importance for the study of stress (about which more below).

Gigerenzer has also taken the correspondence and optimality point of view, and in accordance with tradition, attacked the coherentist view in several articles (see, for example, Gigerenzer, 1991b). The abstract for this article is instructive:

> Most so-called "errors" in probabilistic reasoning are in fact *not* violations of probability theory. Examples of such "errors" include overconfidence bias, conjunction fallacy, and base-rate neglect. Researchers have relied on a very narrow normative view, and have ignored conceptual distinctions—e.g. single case versus relative frequency—fundamental to probability theory. By recognizing and using these distinctions, however, we can make apparently stable "errors" disappear, reappear, or even invert. I suggest what a reformed understanding of judgments under uncertainty might look like. (p. 83)

See also Gigerenzer, Hoffrage, & Kleinbölting (1991):

> Research on people's confidence in their general knowledge has to date produced two fairly stable effects, many inconsistent results, and no comprehensive theory. We propose such a comprehensive framework, the theory of probabilistic mental models (PMM theory). The theory (a) explains both the overconfidence effect (mean confidence is higher than percentage of answers correct) and the hard-easy effect (overconfidence increases with item difficulty) reported in the literature, and (b) predicts conditions under which both effects appear, disappear, or invert. In addition, (c) it predicts a new phenomenon, the confidence-frequency effect, a systematic difference between a judgment of confidence in a single event (i.e., that any given answer is correct) and a judgment of the

> frequency of correct answers in the long run. Two experiments are reported that support PMM theory by confirming these predictions, and several apparent anomalies reported in the literature are explained and integrated into the present framework. (p. 506)

Massaro and Friedman (1990) also express misgivings regarding the oft-cited failure of human judgment to achieve optimality, that is, to conform to the normative (Bayesian) models (and their coherent cognitive activity) under benign conditions:

> In contrast to our conclusions, the consensus from the research is that normative models are invalid [descriptions of judgment processes]. Previous research has rejected Bayes's theorem in various judgmental situations (Kahneman & Tversky, 1972). . . .
>
> Our impression is that the rejections of the Bayesian model have been premature. The rejection of Bayes's theorem in many experiments has been a rejection of the normative form of the model rather than a psychological form of the model. Predictions have been derived on the basis of the objective rather than the subjective sources of information. Our implementations of the models, on the other hand, allow for subjective values for the various objective sources of information. Consider a test of the Bayesian model in situations in which subjective base rates are assumed to be equal to objective base rates. In these cases, performance falls short of the predictions of Bayes's theorem (Leon & Anderson, 1974). Central to the current theoretical framework, however, is the evaluation stage that transforms the objective source of information into some subjective value. Thus, performance could still fall short of the optimally objective prediction but might still be described by the same optimal algorithm if subjective values are assumed. (p. 248)

In short, despite the many challenges to the "optimality" point of view put forward by the coherentists (see, e.g., Kahneman et al., 1982), those emphasizing the correspondence point of view (Anderson, Gigerenzer) continue to defend the optimality position. Indeed, severe criticism of the coherentist's conclusions regarding the nonoptimality of human decision making is voiced by Ward Edwards, one of the founders of the field; he forcefully urges

"cognitive psychologists to change the dismal message they have been conveying to the non-psychological world for the last 20 years! That message, as received, is that People are No Damned Good" (Edwards, in press).

In short, this issue is a dominant one in the field of J/DM today and one that stress researchers cannot avoid.

Interpersonal Comparisons of Intuitive and Analytical Cognition Within the Correspondence Framework

Hammond and his colleagues (Hammond, Hamm, Grassia, & Pearson, 1987) deliberately set aside the question of "normative" performance or "optimality" and made a *direct* comparison between intuitive and analytical judgment in the *same subject*, rather than comparing various forms of cognitive activity to an absolute standard of rationality. They offered the following reasons for this choice:

> Research in the field of judgment and decision-making often compares the rationality of a person's intuitive judgments under uncertainty with analytically derived answers produced by a formal model such as Bayes' theorem, a multiple regression equation, or other rules from the conventional probability calculus. . . . Such comparisons are *indirect:* they compare a *person's* intuitive efforts with *person-independent* operations. That is, they compare a person's intuitive processes and judgments with those of an analytically derived rule or equation put forward as a standard of rationality. Indirect comparisons are undeniably important, but they are necessarily restricted in three ways. First, because indirect comparisons evaluate intuition with respect to a standard of rationality, researchers must choose one standard from among the many offered. However, agreement on which standard of rationality is correct has never been achieved. The choice of any standard, therefore, is subject to dispute, and any conclusions that subjects have failed to achieve the standard chosen are sure to be criticized by those who prefer a different standard (as indeed they have been), . . . thus leaving the comparison between intuitive judgments and rationality unresolved.

> Second, indirect comparisons cannot fail to show that analytical cognition is equal or superior to intuitive cognition because analytical models, however chosen, provide the standard to be achieved by persons. If intuition offers an advantage over analysis, as many have argued it does, its putative advantage cannot be demonstrated in indirect comparisons because the analytical model provides a ceiling for performance. Therefore, it is not altogether surprising that populational studies find that few persons' intuitive efforts achieve the standard . . . and none exceed it. . . . .
>
> Third, when indirect comparisons are made, the analytical models are always provided with all the correct (and only the correct) substantive information each model requires, and such models are almost always executed without error—at least in academic journals. In practice, however, the analytical cognition of persons, in contrast to analytical computation by formal models, is vulnerable to substantive failure (insufficient information, incorrect information, incorrect substantive theory) and to procedural failures (incorrect assignment of numbers to the symbols of the equation, computational errors, use of an incorrect model, insufficient time). In short, valuable as indirect comparisons may be, these restrictions prevent them from informing us about the relative efficacy of the intuitive and analytical cognition of people.
>
> Therefore, *direct* comparisons between a person's use of intuition and the same person's use of analysis are also needed. Direct comparisons will inform us about the *relative efficacy* of these modes of cognition in terms of empirical achievement or correctness. Comparisons of relative efficacy, however, require the presence of an empirical criterion with which judgments are compared, rather than a standard of rationality. When a criterion is available, direct comparisons enable us to address the age-old question: does a person's intuitive or analytical cognition produce more empirically accurate answers? (pp. 753-754)

These researchers found that task conditions strongly influenced the relative efficacy of each of these modes of cognition. In the final section of this report ("New Approach") we build on this conclusion.

Implications for Judgment Under Stress

Despite the division between an emphasis on coherence and correspondence, researchers in both areas are in apparent agreement on one conclusion: Both the rationality and empirical performance of human judgment falls short of what is expected—or hoped—of it, although the coherence researchers are far more persuaded of this "gloomy conclusion" than are the performance researchers.

These gloomy conclusions produced by academic researchers are taken seriously by those interested in J/DM in the work place, particularly if the workplace is aviation. Telfer (1989), for example, writing in a book entitled Aviation Psychology states:

> There are two clear options to identify pilot judgment and decision-making in a theoretical context; the choice aligning with one's interpretation of man's ability as a decision-maker. One view is that people are inherently subjective and fallible in their decisions, the other view being far more positive. This dichotomy in the literature is not new (see, for example, Williams and Hopkins, 1958) and enables the researcher to choose between two polarized perspectives, the pessimists and the optimists (Jungermann, 1983).
>
> The Pessimists
>
> This group emphasizes the human's characteristic inability to remain detached and objective when decisions are needed. An in-flight decision is subject to quite a range of biases which can affect the retrieval of relevant experience and, thus, the pilot's ability to implement an appropriate response (Kahneman et al., 1982). For example, consider the pilot who suffered a reprimand after previously seeking assistance when unexpected IFR conditions were encountered; or the captain rebuked by management for overnighting a passenger complement instead of completing the scheduled flight. The bias against at least one decision option in each case is apparent; against calling for assistance in the first example, and against declaring a no-go in the second.

Just as we suffer perceptual illusions when runways slope or lighting is deceptive, so there can be "editing" defects as we process information. The true range of possibilities may not be considered because of personal consequences as a result of prior events. The result can be an inconsistency in the selection of preferences because of a deficiency in the pilots' perceptions (Tversky & Kahneman, 1974).

Although rational behavior is the ideal, pilot behavior under stressful conditions may be far from ideal. A more hazardous possibility could be defensive avoidance, a coping mechanism by means of which pilots procrastinate, hoping to shift the responsibility, rationalize, or adopt a pattern of selective inattention (Mann & Janis, 1982).

> A representative of the pessimistic school warned: Let us recognize that descriptive theories of decision-making (e.g. Janis & Mann, 1977; Kahneman & Tversky, 1979; Ajzen & Fishbein, 1980) cannot help us out. Their aim is to systematize "how people do it" given various conditions of informedness, time pressure, importance of consequences and the like. But in such theories, too, problem representations are taken as given in advance . . . (Vlek, 1984).

In flying, especially, time is a vital constraint upon decision-making, and only an optimist would assume prior warning of problems. There are, however, decision theorists whose viewpoint could be considered more optimistic. It is to this group that we now turn.

The Optimists

One can take a more positive view of pilots' decision-making ability if the sampling period is extended. Take a series of pilot judgments over the period of a cross-country flight, and both a continuity as well as minor compensations will become apparent (Hogarth, 1981). Separate judgments, examined in isolation, will have limitations but adjustments for these will come with later judgments. From this argument one can conclude that there is a type of meta-rationality demonstrable in the series of judgments which contribute to a successful flight.

A second supporting argument is based upon what could be termed the "cognitive cost" of each judgment to a busy pilot. A sterile analysis of the judgment may point to a gap between the actual and the ideal, but when the competing priorities (control of the aircraft, radio communication, navigation) and the

> available time are considered, the result would have to be regarded as extremely rational (Payne, 1982).
>
> Finally, there is an argument derived from structure. When they make judgments, pilots may be responding to an internal structural representation of the problem (Phillips, 1983). This theory parallels a conceptualization of the manner in which pilots attain flying skills by forming an accurate mental blueprint of what is required in various situations. A similar structure of the essential judgments necessary for safe flight can provide a highly optimistic basis upon which to predict pilot decision-making. (pp. 154-156)

Telfer goes on to describe remedies for the possibility of errors of judgment, about which more below.

The negative conclusions described by Telfer are now so widespread in the field of J/DM (see Christensen-Szalanski & Beach, 1984) that theory and research must take account of them, independently of whether stressful conditions exist or not. That is, stress researchers must now take into consideration the threats to successful cognitive activity so often described by "pessimistic" J/DM researchers in *perfectly benign circumstances* (see, e.g., Kahneman, Slovic, & Tversky, 1982; see also the numerous references to "biases" that are given considerable space in the current textbooks in the J/DM literature; see also the numerous references in non-J/DM literature). For if one accepts the current negative conclusions as true, then the question arises as to where the baseline for evaluating rational or empirically successful performance should lie. And since stress—however defined—is expected to have deleterious effects on cognition, the question arises: What performance level should be taken as "normal," if standards of rationality or empirical performance are not expected to be met even under benign conditions? If we already know that J/DM will be less than optimal, what level of

performance should then be chosen as a baseline against which the effects of stress are to be measured? Exactly what "normative" behavior should be expected from our subjects under benign conditions? If normative standards are irrelevant, must there only be relative standards? How should these be chosen?

If standards of rationality and standards of empirical performance are not expected to be met under benign conditions then between-group comparisons between the performance of experimental (stressed) groups and control (benign circumstances) groups will have to be relied upon. And this indeed is the general methodology that has been employed in (mainly) human factors research. Unfortunately, this approach has neither been as theoretically nor empirically productive as hoped or expected (see comments above by human factors researchers; see also methodology section of this report). Whether the approach taken by Anderson (1990) and/or Hammond et al. (1987) will be more productive in the context of stress research remains to be seen.

Improving Rationality and Performance

Telfer (1989) offers the following description of pilot judgment training (PJT) with which he was familiar:

> As implemented, PJT remains essentially as it was conceptualized by Jensen and designed by ERAU [Embry-Riddle Aeronautical University]. Originally, there was a manual for the student and another for the instructor.
>
> The student manual introduces judgment and the subjects about which judgments are made. The nature of a chain of poor judgments is described, emphasizing the need to break the chain for disaster to be averted. The manual focuses upon recognition of the chain and means of breaking it. By means of

examples, the student is introduced to the mental processes necessary for safe flight.

This section typifies the PJT approach which requires the student to complete exercises in order to gain familiarity with the content. The process of PJT, in both the reading of the manual and in the in-flight components of instruction, is integrated with the content. The program is one of doing and analyses, rather than reading or imitation.

The hazardous thoughts or attitudes are introduced and typified by a brief, easily-learnt phrase. To demonstrate susceptibility, students are asked to complete a forced-choice inventory of situations. The student is able to score the answers, but the results are kept confidential unless the student chooses to discuss them with the instructor. Antidotes to the hazardous attitudes are then presented and learned by the student. Practice is then given for the student to recognize hazardous attitudes and to provide antidotes.

A major section of the student manual deals with stress, its effects on information processing, its origins, recognition, and relieving it. This section deals with life stress as well as that which occurs in flight. The section concludes with a mnemonic, "I'm Safe," which provides a checklist for physical and mental well-being.

The instructor manual assumes a knowledge of the student manual, highlighting the instructor's role and providing materials for guidance. The emphasis is on instructor autonomy in the choice and scheduling of appropriate activities, so that manual tends to be descriptive rather than prescriptive. Guidance is given in the choice of objectives and ways of teaching good judgment. Some examples of scenarios are provided, with further examples of exercises which can be integrated into the private pilot training syllabus. To encourage instructors to develop their own program, a lesson register blank is provided.

To ensure a professional depth of treatment, the stress section in the instructor manual intentionally goes into greater detail than that of the student manual. Recent developments in PJT have been towards greater specialization in the manuals. The Department of Transport and Communication in Australia has in press four PJT manuals: one for student pilots, another for private pilots having an area restriction; a third for unrestricted (cross-country) private pilots, and a fourth for instructors. The USA now has six manuals: student, instructor, helicopter, commercial, Instrument Flight Rated Pilots, and

> Cockpit Resource Management (CRM). There is evidence of wider interest in airline application of PJT to a form of flight crew co-ordination or CRM.
>
> In their application of PJT to instrument flight Jensen and Adrion (1985) used Roscoe's tripartite classification of pilot activities:
>
> 1. Procedural, e.g., managing the powerplant, fuel, navigation, communication
> 2. Perceptual Motor, e.g., control, judging distance, speed; and
> 3. Decisional activities, e.g., self-assessment of capabilities, priorities, hazards.
>
> Jensen and Adrion pointed out that 56% of the non-fatal, pilot-caused accidents result from defective perceptual motor activities. 58% of the fatal, pilot-caused accidents result from decisional activities. This highlights the importance of PJT, and the cognitive judgment commonly termed "keeping ahead of the aircraft". (pp. 161-162)

In general, the results of the studies Telfer describes indicate positive results, although the training procedures themselves seem not to have been guided at all by the academic (largely heuristics and biases) research he cited as a basis for "pessimism." Indeed, there is no further reference whatever in the remainder of Telfer's article to the current body of J/DM literature reviewed here, or contained in J/DM textbooks. Examination of the studies of J/DM that are cited suggests that, although researchers in the field of aviation obviously read at least some of the academic J/DM literature, when they discuss pilot judgment they are much more likely to depend upon research that involves pilots and airplanes.

Telfer is not entirely convinced of the utility of PJT, and notes four problems: "a potential Hawthorne effect, the use of flight training establishments as a source of subjects, the manuals, and testing procedures" (p. 168).

As noted above Telfer (1989) and his colleagues in aviation psychology have taken the "gloomy conclusions" mentioned here as a point of departure for examining pilot judgment under benign and stressful conditions, and examined efforts to improve pilot judgment. Nisbett and his colleagues have also taken note of the negative conclusions about human judgment, particularly as it relates to statistical reasoning and reasoning about economic matters such as "cost-benefit rules of choice" (Larrick, Morgan, & Nisbett, 1990, p. 362), and have undertaken a number of studies to determine whether such reasoning can be improved (see Larrick et al., 1990; Lehman, Lempert, & Nisbett, 1988; Lehman & Nisbett, 1990; Nisbett, Fong, Lehman, & Cheng, 1987). In the abstract of their 1987 article in *Science*, Nisbett, Fong, Lehman, and Cheng describe the pessimistic conclusions:

> Twentieth-century psychologists have been pessimistic about teaching reasoning, prevailing opinion suggesting that people may possess only domain-specific rules, rather than abstract rules; this would mean that training a rule in one domain would not produce generalization to other domains. Alternatively, it was thought that people might possess abstract rules (such as logical ones) but that these are induced developmentally through self-discovery methods and cannot be trained. (p. 625)

But they reach different conclusions:

> Research suggests a much more optimistic view: even brief formal training in inferential rules may enhance their use for reasoning about everyday life events. Previous theorists may have been mistaken about trainability, in part because they misidentified the kind of rules that people use naturally. (p. 625)

An article in the *American Psychologist* by Lehman, Lempert, and Nisbett (1988) reiterated this argument. And in a later article (Lehman & Nisbett, 1990) similar conclusions are drawn:

> The effects of undergraduate training in the natural sciences, humanities, and social sciences on inductive reasoning requiring the use of statistical and methodological principles and on reasoning about problems in conditional logic were examined. Social science training produced large effects on statistical and methodological reasoning, whereas natural science and humanities training produced smaller, but still marginally significant, effects. Natural science and humanities training produced large effects on ability to reason about problems in conditional logic, whereas social science training did not. The improvement in conditional reasoning among natural science students appears to be due, in large part, to knowledge gained in mathematics courses. The results indicate that inferential rule systems, as taught by various fields, can affect reasoning about a wide range of problems. The findings lend support to a version of the premodern *formal discipline* hypothesis. (p. 962)

In a further article in which cost-benefit choices were studied, Larrick, Morgan, and Nisbett (1990) stated:

> Our research shows that people can apply the cost-benefit rules of microeconomic theory to their everyday decisions. Two populations were examined: (a) people who had previously received extensive formal training in the rules and (b) naive subjects who were randomly assigned to receive brief training in the rules. Training affected reasoning and reported behavior in both populations. The results indicate that extremely general rules govern choices across a wide range of domains and that use of the cost benefit rules can be improved through training. (p. 362)

In short, although Nisbett and his colleagues do not discuss the affects of stress on judgment and/or reasoning (nor are their articles cited in J/DM textbooks), the implications of their work are clear: Training in reasoning and/or judgment can result in the rejection of the negative conclusions so often stated. Whether such training can offset the putative negative effects of stressful conditions remains an open question.

Mood

Mood and emotion are often mentioned in relation to J/DM primarily by clinical/social psychologists but they are seldom systematically related to this topic. Taylor's (1991) article is an exception; she notes that "negative events are more likely than positive or neutral ones to elicit causal reasoning" (p. 73). Taylor's "negative events" generally refer to events negatively impacting mood, or self-esteem, rather than say, an emergency on the flight deck, in the operating room of a nuclear power plant, or even in the automobile. Nevertheless, the statement above carries implications for the present topic. For example, after reviewing several studies that bear on the above statement, Taylor concludes that "there is some evidence that the search for a causal explanation for negative events is not merely a response to the need to predict and control that event and similar events in the future, but also to explain away the event in a manner that has few lasting implications" (p. 73). In other words, "What happened? I didn't do it!" That explanation may not be very different from a lay person's, but the initial observation that "negative events are more likely than positive or neutral ones to elicit causal reasoning" calls our attention to the role of negative impacts as events that evoke thought, that is, stimulate the search for an explanation in analytical terms. Positive events possibly evoke complacency—perhaps undeserved.

Taylor's review also includes references to work on "emergency" situations and judgment and decision making as well. For example, "Thus, a strong rapid response to negative events,

coupled with a strong and rapid diminution of the impact of those events, may be most effective for the organism in both the short term and the long term" (p. 79). Although the statement hardly seems debatable, Taylor goes on to suggest that "the initial response may enable the organism to overcome *positively biased* [italics added] thought processes to *deal effectively with the emergency* [italics added]" (p. 79). Thus Taylor introduces the idea that negative events induce "positively biased thought processes," presumably a danger to the organism. Nevertheless "muting . . . the impact of the negative event may be essential for the restoration of positive biases that appear to facilitate effective functioning in nonthreatening environments" (p. 79). Although this argument may be difficult to unravel, Taylor is pointing to the possible impact of stressful events on cognitive activity in a manner not ordinarily considered.

Taylor also suggests that there is an "offsetting response to arousal, which occurs automatically as a compensatory process that reverses its [arousal] effects" (p. 72). Could that be interpreted to mean that operators at a control site notice a negative event that occurs briefly, say, but fail to respond because of a compensatory response ("it didn't happen")? Although Taylor's remarks are often difficult to relate directly to J/DM, there is clearly an effort here to relate the effects of (stressful?) negative events to J/DM.

Mano (1990a) notes that in his laboratory study of arousal and J/DM, "steeper penalties lead to lower levels of performance" and suggests that "emphasis on higher threat and possible negative

consequences of failure generated by the anticipated penalty was maladaptive and hindered performance" (p. 160).

Mano also noted that Janis and Mann (1977) had found "similar stress-generated dysfunctional behavior patterns . . . in situations calling for emergency situations when people believed that there was limited time to deliberate." Finally, he observed that "the processes and outcomes of deciding under stress require further theoretical and empirical attention" (p. 160). But Mano acknowledged that "stress . . . is only one explanation of the lower performance of subjects facing steeper penalties." (p. 160) and called for further research to eliminate alternative plausible explanations for his results, and urged, in particular, greater consideration of task conditions, a conclusion almost all investigators reach.

In a recent report, Mano (1990b) notes: "Considerable evidence suggests that variations in the decision maker's emotional state can have a powerful influence on the processes and outcomes of decision making" (p. xx). But Mano points out that "although affect is generally viewed as a multi-dimensional construct (Russell, 1980) previous research linking affect and decision making has investigated the effects of feelings described only on the dimension of pleasure-displeasure (i.e., good and bad mood)" (p. xx). Mano undertook to investigate the effects of both the good-bad mood dimension and the effects of arousal; that is, to examine "the extent to which affect's two primary dimensions, pleasantness and arousal, jointly shape the process of judgment" (p. xx). By combining the study of mood and arousal Mano joins the mood or affect, that is, the *distress* literature

with the *stress* literature, for arousal has been associated with the study of stress for decades—both as an independent and dependent variable (see Bibliography Appendix).

Gilligan and Bower (1984) investigated the effect of emotional arousal on cognition through the use of hypnosis and report that:

> We have been interested in experimentally testing . . . how emotional mood states might influence such cognitive processes as learning, memory, perception, and judgments. The following discussion of our research in this area is divided into four sections. First, we describe our general procedure of using hypnosis to induce and sustain emotional states in experimental subjects. We then summarize our experiments according to four major results: (1) mood selectively biases the recall of affectively toned material; (2) mood enhances the learning of mood-congruent material; (3) the intensity of a mood affects learning differently, depending on the particular mood and the type of materials used; and (4) emotional states can bias many cognitive processes, such as interpretations, fantasies, projections, free associations, personal forecasts, and social judgments. (p. 547)

In their article there is far more emphasis on the effects of emotions, particularly mood, on memory, rather than on judgment. The only reference to contemporary research in J/DM is to the "availability heuristic" (p. 575).

Nevertheless, the authors make strong claims about the effects of mood on social judgments:

> *Summary.* The experiments reviewed indicate that an emotion can have a surprisingly strong influence on how someone thinks and acts in his social world. In one study, mood was shown to bias free association, thematic storytelling, and personality descriptions of familiar people. In another, subjects seemed to be optimistic or pessimistic about the future—whether it involved personal or national affairs—according to their transient mood state. In a third study, whether a person felt socially successful or rejected biased self-observations of his or her socially positive or negative behaviors. Emotion thus seems to be inextricably related to how we perceive and think, influencing them at every turn. Indeed, results reported throughout this chapter suggest that emotion is often a central component of

> cognitive processes in general, and thus that a comprehensive theory of cognition should address it. (pp. 568-569)

See Appendix for articles on stress and J/DM and related performance that have not received detailed attention in the foregoing text of the present chapter.

In sum, the effect of mood and other emotions on J/DM has received almost no attention from J/DM researchers. Nor do there seem to be any major efforts being made to study these potential influences, despite their obvious occurrence in daily life and in the workplace. The clinical-social literature described in Part I places considerable emphasis on these matters, but the work is so far removed from current J/DM research it is irrelevant to the latter.

Summary

Theory and research in the field of J/DM continues to revolve around two main topics related to stress: (a) cognitive competence in terms of rationality and coherence, and (b) cognitive competence in terms of empirical achievement or performance. The past decade has produced a preponderance of gloomy conclusions regarding both. However, numerous, and often vigorous, doubts have been raised about the generality and applicability of these conclusions.

This dispute has recently been cast in the context of optimality in strong articles by Schoemaker (1991) and also by Anderson (1991b) in the journal, *The Behavioral and Brain Sciences*. The thoughtful commentaries by numerous authors that follow give even more significance to these important articles. The discussion of optimal behavior carries considerable importance for the study of the effects of stress on human judgment, for without a clear conception of optimal

behavior, and what constitutes a proper normative model, stress researchers cannot evaluate departures therefrom (except in extreme cases). They must therefore resort to experimental/control group comparisons of differences in performance, without knowing how far from optimality the performance is. Massaro and Friedman (1990) discuss this issue at a deeper level in connection with what they call the "identifiability problem."

The uncertainty that remains about the quality of human judgment is regrettable, for many researchers begin their studies on the uncritical premise that the gloomy conclusions (Telfer's 1989 pessimistic conclusions) are true and simply seek to confirm them in their work. Others take the opposite position. Worse still, Brehmer (1991) points out that the results are *predetermined* by the different methodologies employed. And he argues that the problem *must* be solved: "This gap must be bridged; psychology cannot, in the long run, live with two such different conceptions of man" (p. 2).

3. Stress and J/DM:

Contributions from Textbooks, Anthologies, and Reviews

Judgment and Decision Making

In what follows, examples are presented of the treatment given to the effects of stress on judgment in various current textbooks.

Textbooks

- Baron, J. (1988). *Thinking and deciding.* Cambridge: Cambridge University Press.

Baron discusses the topic of "rationality and emotion" (pp. 36-37) at length but only in very general terms. Similarly with "rationality and belief" (p. 37ff). And although stress is not indexed by Baron, it receives a subheading under which he states:

> Janis and Mann (1977) proposed that the quality of decision making is affected by "stress,"which occurs when it is difficult for the decision maker to see how to avoid extremely negative outcomes. Excessive stress leads to a state of "hypervigilance," in which the decision maker considers one option after another, with little search for evidence. When the decision maker does seek evidence, the search is unsystematic, and the most useful evidence is often overlooked. (p. 280)

He also refers briefly to mood by referring to Johnson and Tversky (1983) and notes that "availability can be affected by mood" (p. 212).

In short, Baron does not provide a detailed account of the effect of stress or emotion on J/DM, other than a general reference to what Janis and Mann (1977) "proposed."

- Bazerman, M. H. (1990). *Judgment in managerial decision making* (2nd ed.). New York: Wiley.

Neither stress nor mood are indexed by Bazerman. He notes (p. 9) that emotion is a "viable topic of inquiry" for judgment research but indicates that this topic will not be addressed.

- Dawes, R. M. (1988). *Rational choice in an uncertain world.* San Diego: Harcourt, Brace, Jovanovich.

Neither stress, nor mood, nor emotion are indexed or discussed by Dawes.

- Hogarth, R. M. (1987). *Judgement and choice: The psychology of decision* (2nd ed.). Chichester: Wiley.

Neither stress nor mood is indexed in Hogarth. He does note that "Emotional factors such as anxiety, . . . fear of potential outcomes of one's actions, can also cause people to block out relevant arguments, overemphasize arguments in favour of preferred alternatives, fail to search for new alternatives, and even psychologically prepare themselves for negative consequences of their decisions" (p. 111), although no evidence is provided for the truth of these assertions.

Hogarth also comments on time pressure: "Time pressure tends to induce greater use of noncompensatory strategies as people seek to find means to process information under greater constraints" (p. 81). He bases this conclusion on Wright (1974); but as noted above, these results were not replicated by Svenson, Edland, and Slovic (1990).

Hogarth also comments on what Yates would call "task stress": "Complexity induced by time pressure, information overload, distractions lead to reduced consistency of judgement" (p. 220).

Again, however, no support for these assertions is provided. And he makes a brief comment about emotion: "Emotional stress reduces the care with which people select and process information" (p. 220).

- Margolis, H. (1987). *Patterns, thinking, and cognition: A theory of judgment.* Chicago, IL: University of Chicago Press.

Margolis does not discuss stress, mood, or emotion.

- Rachlin, H. (1989). *Judgment, decision, and choice.* New York: W. H. Freeman.

Rachlin does not consider stress, mood, or emotion.

- Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty.* Reading, MA: Addison-Wesley.

Although not to be considered "current" this textbook is mentioned here because it was one of the most widely used books in the introduction and development of the field of J/DM. There were many books to follow in this genre, but, like this one, they do not consider the effects of stress.

- Russo, J. D., & Schoemaker, P. J. H. (1989). *Decision traps: Ten barriers to brilliant decision-making and how to overcome them.* New York: Doubleday.

The only reference in this text to stress occurs when the authors indicate that, after "analyzing . . . [certain] catastrophes, researchers discovered common elements that—though apparently innocent—seemed to lead toward tragedy," among which they list "high stress," described as follows: "The importance of the decision, its complexity, and tight deadline put group members under great pressure" (p. 148). They observe that "cohesiveness, insulation, and

stress generally led the groups [studied] to reach consensus too quickly" and they "focused almost exclusively on information that confirmed their opinions" (p. 148). The authors rely on work by Janis for these conclusions. Given the type of textbook prepared by Russo and Schoemaker it apparently was unnecessary to document these conclusions in detail.

- Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice-Hall.

Of the several textbooks examined, Yates presents the most material (two pages) on stress and J/DM. Three features of Yates's treatment should be noted: (a) He carefully draws the same distinction between "ambient stress" and "task stress" that Hancock and Warm (1989) and other human factors researchers have made between "exogenous" and "endogenous" task conditions inducing stress); (b) he relies on the Yerkes-Dodson "law" that has largely been discarded by human factors stress researchers; and (c) he emphasizes the "narrowing" hypothesis but is not uncritical of its application.

Because Yates's book is among the most recent (1990), and because he gives the greatest attention to the topic of stress and judgment, it is worth noting the dates of the publications he must rely on: J. Anderson (1976), Ben Zur and Breznitz (1981), Easterbrook (1959), Janis and Mann (1977), Lazarus (1966), Rothstein (1986), and Wright (1974). (See Part I for comments on these studies.) Only the Ben Zur and Breznitz and Rothstein studies were carried out within the past decade; moreover, the studies cited are only tangentially related to one another, and the study by Wright (1974), upon which

much of the discussion depends, has failed to be replicated (see Svenson et al., 1990). In short, when the author of a textbook does make a responsible attempt to describe the known effects of stress on judgment, s/he must rely on material largely 20-30 years old, and rest his/her discussion on studies of dubious validity because of lack of confirmation. Demonstrably sound, replicated results are not available.

In sum, current textbooks on J/DM devote little space—many devote none—to either the effects of stress, emotion, or mood on J/DM. No textbook devotes as much as a chapter to these topics and many do not even index these terms. Thus, the writers of textbooks have found what is reported in Part I; very little *psychological* research or theory has been devoted to the effects of stress on J/DM. But it must be said that the textbooks that purport to describe the field of judgment and decision making ignore a great deal of research that has been carried out by human factors researchers.

Anthologies/Reviews

• Anderson, B. F., Deane, D. H., Hammond, K. R., McClelland, G. H., & Shanteau, J. C. (1981). *Concepts in judgment and decision research: Definitions, sources, interrelations, coments.* New York: Praeger.

This book serves as a glossary for the field of J/DM. Stress is not indexed.

• Arkes, H. R., & Hammond, K. R. (Eds.). (1986). *Judgment and decision making: An interdisciplinary reader.* Cambridge: Cambridge University Press.

Intended to be an interdisciplinary reader, this book contains 43 chapters under 11 topics. Stress is not indexed, although "emotion" is, in a chapter by Herbert Simon. Three pages are devoted to a discussion of "intuition and emotion," "emotion and attention," and "emotion in education." These remarks generally concern the role of emotion in learning rather than J/DM, and generally describe the positive contributions of emotion in that process.

- Brehmer, B., & Joyce, C. R. B. (Eds.). (1988). *Human judgment: The SJT view.* North-Holland: Elsevier.

This book provides an overview of social judgment theory (SJT). Stress, emotion, and mood are not indexed.

- Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). *Human judgment and decision making: Theories, methods, and procedures.* New York: Hemisphere/Praeger.

This book provides the first overview of the contemporary field of J/DM. Six approaches were reviewed and compared with one another. Neither stress, nor mood, nor emotion were indexed, and there is no discussion of these topics.

- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

There is no reference to stress, mood, or emotion in this anthology, which presents the first organization of theory and research that is the foundation of the "heuristics and biases" approach.

Human Factors

Books

Four recent books that belong to the human factors literature rather than the J/DM literature are described below because they are very likely to be influential, and because they are not mentioned in the J/DM texts.

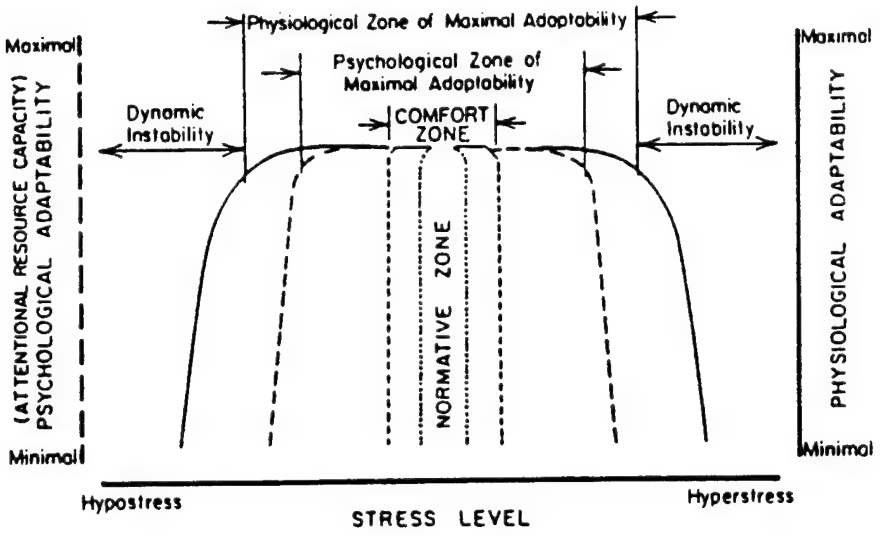
- Hancock, P. A. (Ed.). (1987). *Human factors psychology.* North-Holland: Elsevier.

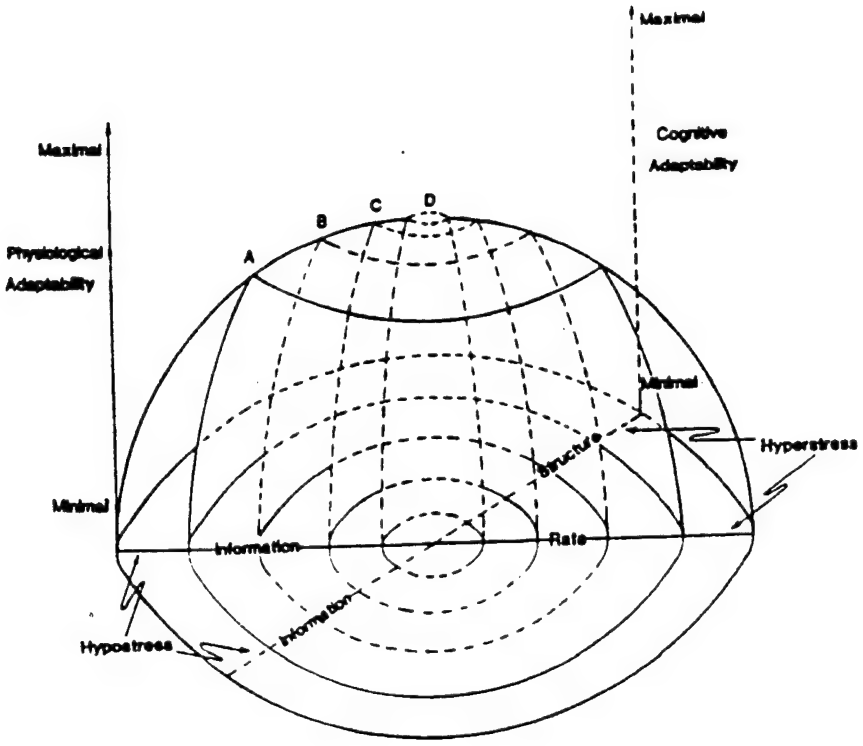
Hancock's edited volume also provides some link to the J/DM literature. He provides numerous references and descriptions of studies of cognitive activity in relation to various environments in which stresses of various types occur.

"Stress" receives 12 entries in the index; "work load" receives six additional entries; none of the references in the bibliography are included in any of the J/DM textbooks in their discussions of stress, thus indicating once more the independence—Reason excepted—of the human factors literature and the J/DM literature. Hancock's book also includes an excellent review chapter by the J/DM researcher, D. Kleinmuntz, that describes the nature of J/DM theory and research, and thus represents a conscious effort by the editor to bridge the gap between the two literatures.

Hancock and Chignell (1987) offer what is perhaps the most detailed theoretical treatment of stress apt to be found in any literature (see pp. 309-319). They describe the "origins of investigations into stress effects" (p. 309) beginning with reference to Cannon's concept of "homeostasis," then observe that "perhaps the

most influential studies of [stress effects on performance] . . . emerged from the Applied Psychology Unit in Cambridge, England, . . . the staff of which included Bartlett, Broadbent, Poulton, Welford" (p. 309). Hancock and Chignell then observe that the "fruitful integration with arousal theory" that invigorated stress research during the post-war years, "has begun to fail as a descriptive account of empirical findings, while its theoretical foundations have always been somewhat suspect" (pp. 309-310). They further note that "in the vacuum that has followed the conceptual disintegration, a number of novel and synthetic alternatives have been postulated." The general ideas developed by Hancock and Chignell can be seen in their two diagrams (pp. 314 and 316):





● Reason, J. (1990). *Human error.* Cambridge: Cambridge University Press.

Although Reason's book was not intended to be a textbook for cognitive psychologists, it contains a great deal of material relevant to judgment and decision making. And because it is obvious that Reason has read widely in this area, as well as in the field of systems analysis and design, his observations are important.

"Stress" is indexed and receives three references. In his discussion of "stress and knowledge-based errors" Reason, after describing Brehmer's work (Brehmer, 1987; Brehmer, Allard, & Lind, 1983) on subjects acting as "fire chiefs" controlling a forest fire, notes

that Brehmer's results indicate that "the tendency to overcontrol increases as a function of stress" (p. 92), a finding that Reason finds consistent with that of Doerner's (1987) conclusions.

Reason also discusses stress in relation to "training errors" (p. 245). He suggests that there are "negative aspects of training errors [that] have to do . . . with the trainee's motivation and self-appraisal" and that the affective component of feedback regarding performance is important; for example, "errors . . . can lead to self-blame and additional stress." And "stress and anxiety increase the cognitive load upon the trainee, which in turn promotes the occurrence of further errors" (p. 245). He does not cite evidence.

Reason discusses what he calls the *stress-vulnerability* hypothesis in some detail:

> Observations taken from a wide range of samples, in which people were exposed to different stresses (nurses under training, women facing breast surgery, students preparing for important examinations), support Broadbent's *stress-vulnerability hypothesis:* namely that relatively high levels of cognitive failure in normal everyday life are associated with increased vulnerability to externally imposed stresses. Whatever governs general proneness to everyday slips and lapses also appears to contribute to stress vulnerability. This factor eludes capture by most simple laboratory measures, [see methodology section for more on this topic] but it seems to relate closely to the deployment of limited attentional resources in the face of competing task demands. The evidence so far assembled suggests that it is not so much that stress induces a high rate of cognitive failure, but that certain styles of cognitive management can lead to *both* absent-mindedness *and* to the inappropriate matching of coping strategies to stressful situations. (p. 15)

Reason bases these provocative conclusions on several studies from the psychological literature—largely individual differences.

Reason indexes and discusses (briefly) the availability and representativeness heuristics. But he is influenced far more by design

engineers such as Rasmussen. Thus, for example, Tversky is cited six times, but Rasmussen is cited 21 times, and Norman is cited 27 times. His analysis of human errors depends heavily on Rasmussen's skills, rules, and knowledge framework that has also heavily influenced the engineering approach often published in journals by the Institute of Electrical and Electronics Engineers. Judging from literature citations, psychologists rarely read this literature; it is entirely absent from J/DM textbooks.

Thus Reason goes further toward bridging the gap between the J/DM literature and the human factors literature than any other author.

• Senders, J. W., & Moray, N. P. (1991). *Human error: Cause, prediction, and reduction: Analysis and synthesis.* Hillsdale, NJ: Erlbaum.

The material for this book grew from a series of conferences beginning in 1980. The authors indicate that it represents the conclusions drawn by the 22 participants regarding the state of knowledge regarding human error. Although the topic of stress receives no separate detailed discussion, it does permeate the text in a variety of ways and at a number of points, and therefore, there is much to be gained for the reader interested in the topic from reading the book.

In addition, this book emphasizes social or organizational sources of stress and/or error more than is customary. For example, Wreathall, one of the participants in the conference, observes:

> We have had several major accidents that result from complex error mechanisms that are apparently organizational or social in

> origin. Examples include Chernobyl, Kings Cross, Herald of Free Enterprise, Bhopal, and Challenger. (p. 134)

This conceptual framework was further developed by Hancock and Warm (1989).

- Wickens, C. D. (in press). *Engineering psychology and human performance* (2nd ed.). New York: Harper Collins.

Wickens devotes an entire chapter to stress and human error and thus provides a major source of information for this topic. Ninety-seven items are included in the bibliography; the coverage is broad and up-to-date. There are no references to the J/DM research programs to be described below, however, presumably because these have little to say about stress and errors of judgment. The material in the chapter is therefore drawn primarily from the human factors literature. Because Wickens's chapter is part of a textbook, it is integrative and explanatory in tone and purpose, and thus serves a valuable function for the present review.

Wickens begins by offering definitions of stress. Surprisingly, these do not include the usual references to endogenous and exogenous factors; the emphasis is, however, on "*stressors* . . . [that influence] information processing and cognition which are not inherent in the *content* [italics added] of that information itself" (e.g., noise, vibration, etc.) Wickens thus omits references to changes in the *formal* properties of the task itself (aside from content) that putatively induce stress. He states that such (exogenous) stressor "have three manifestations": (a) they affect phenomenological experience; (b) physiological changes occur; and (c) the efficiency of

information processing is affected—although Wickens is careful to point out that "such effects . . . need not always degrade performance." Wickens suggests that "many of the effects are mediated by arousal, and these will represent the major focus of the following discussion."

The remaining discussion does indeed follow the promise. Wickens is aware of the Yerkes-Dodson Law; he points out that "it is somewhat difficult to know, a priori, where the optimum level of arousal is for a particular task, and hence, whether the introduction of a stressor will lead to an initial increase or decrease in task performance. This shortcoming makes the law inadequate as a predictive model for stress effects." Nevertheless Wickens does find this law "useful in understanding stress effects." He also finds it important to note that the inverted-U shaped curve is different for simple and complex tasks, a feature that allows for an explanation of the differences in effects of stress on unskilled and skilled persons (effects for the latter are less).

Wickens refers to Hockey's research (1984, 1986) that offers "signatures" for various stressors. This matter requires too much detail to be presented here, but certainly stands as one of the major advances in the discussion of this topic. Wickens also notes the equivocal nature of the results of "attentional narrowing," a response that may degrade *or* enhance performance, depending on what is required of the person who must act. (The failure at Three Mile Island is offered as an example of "tunneling" that led to degraded performance.) Long term memory and short term memory are topics included in this discussion.

Wickens discusses the "speed-accuracy tradeoff" that has been observed to occur under stress, and notes that it has been accepted as fact in "many countries [that] have explicitly requested that operators carry out no physical actions for a fixed time following an alarm, while they gain an accurate mental picture of the nature of the malfunction."

Decision making receives a subheading in this chapter but Wickens does not offer a definite conclusion regarding the effects of stress, nor is there any reference to the J/DM research programs discussed here.

Wickens's discussion of "coping with stress" is limited to a few studies; pilot training is mentioned and the positive effects of extensive training are emphasized. As an example, he notes the need for training subjects when under stress, to engage in behavior opposite to that normally required. He then cites the example of losing control of an automobile when driving on ice, where one is required to turn in the direction of the skid. An important example of very rapid learning to do the opposite of what is normally required can be found in the record of UAL 232 (loss of hydraulics) where the Captain (Haynes) reports learning to do just the opposite of what is required when the airplane stalls under normal conditions, a case of rapid perceptual-motor learning on demand under high stress. A form of analytical cognitive change can be seen in the report of Commander Rinn (Captain of the USS Roberts; see Part I of the present monograph) where he reports doing the opposite of what "doctrine" required in order to save a sinking ship.

Wickens concludes his discussion by observing that "it is clear that considerably less is known about stress effects, and the appropriate techniques for their remediation, than about many other aspects of performance."

Following his discussion of stress effects, Wickens turns to a discussion of human error. He emphasizes the receiving of research and theorizing about this topic and "chart[s] the recent developments that have taken place in categorizing human error within a framework that is consistent with the information processing model presented in this book." He then emphasizes the "slips and mistakes" dichotomy introduced by Norman, and relies heavily on the work by Reason and Rasmussen. This material is not described here because it has not been brought closely to bear on the topic of stress and judgment. Nevertheless, it is recommended reading for psychologists in the J/DM field because it will be informative with respect to the types of task environments faced by persons required to exercise their judgment under stress. Information of this sort will be sobering for psychologists who use simple laboratory tasks.

Articles

Simmel, Cerkovnik, and McCarthy (1989) provide results that conflict with the conventional wisdom regarding "life stress" and judgment, but which are compatible with results from J/DM studies regarding incorrect assessment of the seriousness of events:

> The initial stages of our research on the effects of the overassessment or underassessment of the consequences of nonroutine events (a potentially important factor in decision-making by pilots) is described. Critical difficulties in the measurement of life stress are discussed and a partial solution is

> suggested in the form of Life Events Questionnaire (LEQ), an instrument on which individuals report the nature of stressful events, when they happened, and how much they concern them at the time of testing. The results show that the tendency to overassess or underassess the consequences of nonroutine events is consistent within individuals, but is unaffected by life stress levels on the artificial air traffic control simulation used. Analysis of LEQ results showed that high stress subjects had much higher chronic than acute stress scores ($p < 0.001$), although the chronic and acute scores were virtually identical for low stress subjects. (p. 53)

Entin and Serfaty (1990) introduce their work with a comment on "stress-induced cognitive disruptions" that summarizes what they apparently believe to be the current state of knowledge regarding that topic:

> Under stress, particularly high stress, persons may experience such an inordinate amount of cognitive constriction and perseveration that their thought processes are disrupted (Janis & Mann, 1977). Normally, during periods of nonstress, our thinking is essentially rational, logical, and flexible. During periods of stress, thinking is often dominated by worries about the consequences of our actions and by negative self-evaluations. One's normal patterns of organized, logical, and coherent thinking are impaired. Individuals in stressful environments report worrying about possible failure and about their own inadequacies. This, in turn, interferes with thinking about the tasks they have to perform (Spielberger, 1979). High levels of stress disturb individuals' ability to concentrate and the process of selective attention, i.e., the ability to concentrate on specific stimuli while ignoring other irrelevant stimuli. In such stress situations, persons report being distracted both by obsessive thoughts of failure and by external stimuli. Obviously, poor concentration impairs an individual's performance and decisionmaking ability.
>
> Stress-induced memory impairment is most likely due to the deterioration of the ability to transfer information from short-term to long-term memory. The rehearsal process (necessary for short- to long-term transfer) appears to be disrupted under stress, leading to frequent confusion about the sequence of events (Rimm & Somervill, 1977).
>
> Janis and Mann (1977) report the immediate memory span of individuals is reduced and their thinking becomes more simplistic. That is, they cannot deal conceptually with as many

> categories as when they are unstressed. People fail to recognize all the options open to them and fail to use remaining resources to evaluate adequately those alternatives of which they are aware. Under high stress people are likely to search frantically for a solution, persevere in their thinking about a limited number of options, and then stick tightly to a hastily contrived solution that appears to promise immediate relief.
>
> In summary, stress is capable of disrupting a variety of cognitive functions. This implies that decisionmaking, which usually requires a combination of good concentration, flexible thinking, intact memory, and visual imaging, will be significantly affected by stress. (pp. 9-10)

Conclusion

There is little or no information to be gained about stress and J/DM from J/DM textbooks in this field. Books written by human factors researchers are more informative; for example, Hancock and Chignell (1987) provide a schematic for assistance in understanding the process, and Wickens (in press) offers a very useful summary of what has been done in this area. But despite hints of change that appear in a few articles, there remains a wide gap between the J/DM researchers and the human factors researchers in the area of stress.

4. Current Theories and Models of J/DM: Implications for Stress Research

There are a variety of approaches to the study of J/DM that have persisted for several decades. These were referred to in the Introduction as "research programs," inasmuch as they have persisted in the use of certain concepts and methods, each largely in isolation from others. When the proponents of one research program do refer to another, the tone of the reference is likely to be negative (see, e.g., Anderson, 1981, 1982)

In what follows we describe several of these research programs and indicate their potential usefulness for the study of the effects of stress on J/DM.

Formalized Models

Signal Detection Theory

Signal detection theory arose from a growing realization among researchers studying visual and auditory perception that perceptual judgments (e.g., whether a stimulus is present) could not be adequately described by models which assumed an *absolute* threshold of detection. Instead, efforts to determine the minimum stimulus value required for a response showed significant variation, indicating that there are factors in addition to the judge's discriminatory capacity and the magnitude of the stimulus which determine an observer's response. At the same time, psychologists were beginning to develop cognitive process models which identified certain of these nonsensory factors and could explain their influence on perception. Broadbent (1957, 1958, 1971), for example, developed a theory of attention in

which sensory input must first pass through a selective filter and a channel of limited capacity before detection. Variation in the level of attention given to a stimulus, then, could lead an observer to set a different decision criterion for determining whether or not a stimulus is present, which in turn could explain some of the variance in the measured threshold of detection. In this way Bayesian probabilism became formalized in judgment theory.

The resulting theory of signal detection (Green & Swets, 1966/1974; Swets, 1973) partitions human judgment behavior into two independent components: an information processing component which builds internal representations of external events and a decision component that generates responses. The information processing component builds representations of two categorical events (e.g., whether a visual stimulus is present or absent, or whether a defendant is guilty or innocent) which are modeled by overlapping Gaussian probability distributions. The accuracy with which the information processing component can separate the two distributions is a function of the means and standard deviations of the two distributions and is independent of the criteria or decision rules used to generate responses. The decision process compares the internal representation generated on a single trial with one or more internal decision criteria and decision rules are used to generate a judgment about the event (e.g., a simple decision rule might be "state that the event occurred if the evidence exceeds the criterion, state that it didn't occur if the evidence does not exceed the criterion"). There are four possible outcomes for each decision, the relative frequency of

which correspond to different areas under the event probability distributions and bounded by the decision criterion. Only two of these outcomes are independent: the hit rate (the probability of predicting an event given that the event occurs) and the false alarm rate (the probability of predicting an event given that the event does not occur. The other possible outcomes are a correct rejection (the probability of not predicting an event, given that it does not occur, which is the complement of the hit rate) and a miss (the probability of not predicting an event, given that it does occur, which is the complement of the false alarm rate). The relative probabilities of these four outcomes are a function of both the judge's sensitivity and choice of decision criterion.

Significance for the study of stress. A wide variety of important human judgment tasks can be conceptualized in the signal detection theory framework (e.g., polygraphers attempting to discriminate the justly from the unjustly accused; medical doctors making diagnoses based upon medical imaging techniques such as computed tomography; and forecasters trying to predict tomorrow's weather). Because signal detection theory separates the effects of sensitivity and decision criterion, it makes it possible to ascertain whether a stressor independently influences one or the other or both. For example, a weather forecaster under time pressure might become either more or less sensitive to differences in the data being evaluated when using any given decision criterion. The time pressure also may cause a forecaster to shift his or her decision criterion, affecting the trade-off between false alarms and misses. In a recent study which applied

signal detection analysis to forecasters who were predicting severe weather at approaches to Stapleton International Airport in Denver (Harvey, Hammond, Lusk, & Mross, in press; Lusk, in press), for example, increased stress was found to increase the forecasters' accuracy while at the same time shifting their decision criteria in the direction of preferring fewer false alarms. Stress therefore induced the forecasters to change their policy for allocating risk: under stress, it became more important for them to avoid false warnings of dangerous weather (which would require re-routing and delaying air traffic) at the expense of increasing their chance of failing to issue a warning when they should (thereby increasing the risk of a weather-related accident). Thus after a long history of laboratory research, signal detection theory has demonstrated its utility for ascertaining the effects of stressful conditions on judgments under uncertainty in field conditions.

Decision Analysis Models

The decision analysis approach to judgment and decision making arose in the 1950s and 1960s when psychologists discovered the theories of optimal behavior proposed by mathematicians, economists, and statisticians (e.g., Raiffa, 1968, in Textbook section). One such theory is the theory of subjective expected utility (SEU) developed by von Neumann and Morgenstern (1947). They demonstrated mathematically that, if a decision maker's choices follow a certain set of axioms, it is possible to derive expected utilities for each choice (the sum of the subjective value of each potential consequence of the choice multiplied by its probability) such that a

given choice will be preferred to an alternate choice only if its expected utility is higher. SEU theory therefore describes a normative or "rational" model of how a decision maker should combine expected utilities to make optimal choices. Another influential normative theory is Bayes' theorem, which was proposed by Savage (1954) as an optimal mathematical formula for using new data to update prior beliefs about the probability of an uncertain event. Recognizing that these theories represented formal models of individual judgment behavior, psychologists began to test them empirically and suggest revisions. Edwards (1954), for example, demonstrated through a series of experiments on risky choice that SEU theory must take into account the subjective nature of probabilities as well as that of values.

There are two basic ways in which decision analysis models can be used. First, they afford a rational analysis of the required judgment, which yields a normative theory (e.g., SEU theory or Bayes' theorem) of how the decision maker *should* behave in order to make optimal judgments. Second, they can provide a descriptive analysis of human beliefs and values and the manner in which they *are* incorporated into actual judgments. Subjects' judgments may then be compared with those produced by the normative theory; and the comparison thus forms a "yardstick" for determining how close human judgment behavior is to the best it could possibly be. Decision analysis models need not take a position on the cognitive processes by which judgments are derived; their focus is on comparing the rational model with measured behavior.

Measured judgments can deviate from the normative models in one of two ways: by being less *accurate* (following the axioms of the normative model but not optimally) and by being less *coherent* (violating the underlying assumptions of the normative model). Edwards (1968), for example, compared human probabilistic inferences with those prescribed by Bayes' Theorem and reported that, although people's judgments were proportional to the optimal Bayesian calculations, they typically underestimated the optimal responses. Thus, although Edwards' subjects behaved consistently with the normative model, they were less accurate. Other studies have reported probabilistic inferences to be less coherent than they should be, violating the assumptions of the Bayesian model, for example, by failing to take into account relevant base rate information (Kahneman & Tversky, 1973).

When people deviate from the rational model, the typical response of decision analysts is to use decision aids to help them implement the rational model. Raiffa (1968, 1969), for example, developed multiattribute utility theory to extend SEU theory to the case when each choice option has multiple attributes, and it has been used to aid judgments in such applied situations as expert evaluation of water quality and deciding where to locate a new airport (see von Winterfeldt & Edwards, 1986, for a description of the use of decision aids).

In a significant shift of emphasis from optimal model accuracy and coherence John Anderson (1990) has recently proposed a "rational analysis" of human cognition which applies essentially the

same normative/descriptive approach to memory, categorization, causal inference, and problem solving that decision analysts have applied to probabilistic inference and choice. The rational theory in this case argues that human behavior is an optimal response to the environment in which it typically operates; thus cognition is "adaptive." Anderson concludes that cognition is *empirically* accurate, that is, offers judgments that correspond to the correct answer. Moreover, this achievement can be explained by identifying the appropriate rational model and the structure of the environment, without having to postulate internal cognitive mechanisms.

Significance for the study of stress. The major advantage that decision analysis offers for studies of stress and decision making is that, when there is agreement on which normative theory is appropriate, it provides a baseline or yardstick for assessing (optimal model) accuracy and coherence. For example, if one knew how far people deviate from Bayes' Theorem under normal circumstances, one could then determine the effects of stress; for example, whether stress induced persons to become relatively more or less conservative information processors. Or, if people's subjective probabilities and values have been measured within the framework of a multiattribute utility model, one could assess whether the combination of these parameters becomes more or less optimal under stress. The effects of stress on coherence can be examined as well. It is possible, for example, that stress may affect the use or neglect of base rate information in inference. In choice, for example, stress might lead subjects to abandon the complicated multiattribute utility model for a

strategy which uses only part of the available information, such as elimination by aspects (Tversky, 1972). Overall, decision analysis makes it possible to ascertain whether stress affects either (a) the decision to adopt a normative versus some other model or (b) the values of the subjective parameters (e.g., subjective probability or decision weight) if a normative model is adopted. However, at present there is no theory which can predict which parameters would be expected to change in which direction under stress. Decision analysis can, however, provide prescriptions for improving decision making under stress in the form of decision aids based on rational models.

The "log odds" form of Bayes' Theorem (log posterior odds = log prior odds + log likelihood ratio) would prove especially useful for ascertaining the effects of stressful conditions because of its simplicity and its psychological significance. The simplicity of the formulation is obvious; wherein lies its psychological significance? The "log posterior odds" part of the formula corresponds to, and can be compared with, the judgment of the subject regarding the likelihood of an event. The "prior odds" part refers to the "base rate" occurrence of the event. The "likelihood ratio" refers to whatever other evidence may be available to the subject (or investigator). Thus, if the investigator wishes to ascertain the relative impact of base rates versus other information on the subjects' judgment, this formulation makes it easy to do so. And, indeed, various researchers have investigated this question (see, e.g., Kahneman & Tversky), leading to the conclusion that people frequently "ignore" base rates in favor of other evidence.

The question of interest is whether stressful conditions lead to changes in either component of the above equation; that is, does stress lead to *better* or *worse* subjective judgments of the posterior odds of the occurrence of an event? And does either the improvement of decrement in accuracy occur because of change in one component or the other?

Heuristics and Biases

The approach to human judgment known by the popular label "heuristics and biases" arose out of several separate lines of psychological research in the 1950s and 1960s which collectively demonstrated that human judgment is often poor in comparison to normative models. Meehl (1954), for example, in a comparison of predictions made by clinical psychologists and by a simple linear combination of the cues which the clinicians used to make their judgments, demonstrated not only that the statistical model outperformed the human judges but that the judges greatly overestimated their ability to make accurate predictions. At the same time, Edwards (1954) was documenting substantial discrepancies between human inference and optimal Bayesian models, and Simon (1955, 1956) was developing his theory of "bounded rationality," which postulates that individuals do not search for optimal choices but instead, due to time constraints and limited computational capacity, seek a solution which satisfactorily meets their level of aspiration and then terminate their search. Eventually, so many examples of judgmental error or bias had been collected that many researchers felt it was time to abandon normative models such as SEU theory and

Bayes' Theorem as inadequate for descriptive purposes and search for alternate models of human behavior. Judgment researchers turned to the emerging field of cognitive psychology for these new models, inspired by cognitive process models which emphasized limitations in such mental capacities as short-term memory (Miller, 1956) and attention (Broadbent, 1957) and how these limitations shaped more complex cognitive processes such as reasoning (Bruner, Goodnow, & Austin, 1956) and problem solving (Newell & Simon, 1972). Tversky and Kahneman (1974), for example, have proposed several simple cognitive processes, including representativeness, availability, and anchoring and adjustment, which they claim underlie a broad range of human judgments since they reduce the complexity of judgmental tasks and make them more tractable for decision-makers with limited mental resources.

Like decision analysts, researchers in the heuristics and biases tradition are interested in comparing normative models of decision making with empirically derived human judgments. However, they have a different focus: the goal of many studies in this tradition is to specifically search for situations in which judgments can be considered to be in error or biased with respect to the normative model. Once such an error has been demonstrated, it is typically explained by postulating that people employ simple "rules of thumb" or "heuristics" which save time and effort. It is proposed that these heuristics often lead to accuracy because they approximate the appropriate normative model in most situations. However, under certain circumstances they can lead to severe and systematic errors.

It is the heuristic cognitive model, rather than the normative theory, which is used for subsequent explanation and prediction.

Since normative theories are deemed to be inadequate for descriptive purposes, heuristics and biases researchers propose alternate descriptive models which incorporate heuristic rules. Kahneman and Tversky (1979), for example, have developed prospect theory as an alternative to SEU theory. Prospect theory extends SEU theory by incorporating an initial "editing" phase in which various heuristic rules are used to simplify decisions, by postulating that values are assigned, not to overall assets, but to gains or losses with respect to a reference point, and by assuming a generally steeper value function for losses than for gains. The theory can account for a variety of empirical effects which SEU theory cannot, including the tendency to overweight certain outcomes relative to possible outcomes with the same expected value and reversals of preference due to framing effects. Another example of a formalized heuristic theory is the contingent weighting model proposed by Tversky, Sattath, and Slovic (1988) to explain choice between two options. In their model the weight or importance assigned to different input attributes of the choices depend on the nature of the output which is required. The theory can therefore account for response mode effects (e.g., reversals of preference depending on whether the decision maker is asked to choose between two options or to adjust the inputs of the two options so that they are equally matched) which Multiattribute Utility Theory (MAUT) cannot. Many of the proposed cognitive heuristics are not so well formalized as prospect theory and contingent weighting theory.

The "heuristics and biases" replacement for Bayes' Theorem, for example, is to turn to a collection or "toolbox" of heuristics which are often not well defined. Instead of using base rate information to judge probabilities, as Bayes' Theorem requires, people may instead make their judgments by representativeness (Kahneman & Tversky, 1972), or how similar the judged instance is to the relevant category. Or, people may use the information that Bayes' Theorem suggests, but use it in a different way, for example, by employing an anchoring and adjustment heuristic (Slovic & Lichtenstein, 1971) in which they overly focus on one piece of information and adjust it, typically insufficiently, to take other information into account. However, heuristics and biases models at present provide no way of predicting which heuristic will be taken out of the toolbox and applied by decision makers in which situation.

Although heuristics and biases researchers defend their focus on errors and biases as sound research strategy rather than a deliberately pessimistic view of human judgment (e.g., Kahneman, 1991), the list of heuristics and biases is now so long (see, e.g., Kahneman et al., 1982)—and everexpanding—that the field is generally perceived as predicting "gloomy conclusions" for human decision making ability. According to this view the "coherence" of human decision makers is typically very poor, since people adopt heuristics which deviate radically from the procedures prescribed by normative models. Lacking a single rational underlying strategy which they can apply in a variety of situations, people are very susceptible to framing and elicitation effects, and their judgments and choices are easily pushed

around by seemingly subtle differences in the surface structures of problems which are formally identical. Despite this almost complete lack of coherence constancy, it is often claimed that empirical accuracy is usually quite good, that is, that the heuristics approximate the normative models quite well most of the time, despite their flawed logic, and it is only under unusual circumstances that they result in error. For example, frequency estimates based on availability, or "the ease with which examples or instances come to mind" *should* generally be accurate since frequency and memorability should be highly correlated. It is only for cases when memorability and frequency diverge, for example, when a risk is reported more often in the newspaper than its true frequency would dictate (see, e.g., Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978), that relying on the heuristic results in error. However, due to the focus of researchers on locating and studying errors, there is little empirical work on the situations in which heuristics are postulated to work well.

Significance for the study of stress. Heuristics and biases models, when they are formalized, provide an advantage over decision analysis models for studying the effects of stress on cognition because they provide a detailed model of the cognitive process. One can assess the effects of stress upon the parameters in the cognitive model rather than on the rational normative model which people are not likely following anyway. For example, prospect theory might predict that, under stress, people may shift the reference point they use to assess gains and losses. Or, contingent weighting theory might predict that, under stress, people may become more or less likely to

weight inputs depending on their compatibility with the required outputs. When they are not formalized, heuristics and biases models are less helpful for evaluating the effects of stress on cognition, because it is difficult to predict which heuristic will be used in which situation. It is possible, for example, that under stress people would pull an entirely different heuristic out of the cognitive toolbox than they would under normal circumstances, and there is no way of predicting which one.

With respect to coherence, the heuristics and biases approach would predict that, since coherence is so poor even under the best of circumstances, it can only get worse under stress: people employing simplifying rules of thumb are unlikely to switch to highly analytical rational or normative strategies under stress. With respect to empirical accuracy, the heuristics and biases approach might predict either a decrease in accuracy or no change in accuracy, depending on the situation. It could be, for example, that the stressor changes the situation from one in which a heuristic works well, yielding a close match between the judgment and the true environmental situation, to one of those "unusual" situations in which it breaks down. Or, it could be that the judge is already in a situation in which the heuristic performs poorly, so the stressor may have little or no effect. It is difficult to imagine a situation in which a stressor fortuitously changes the circumstances from one in which a heuristic that has been performing poorly becomes more accurate, although it may be possible.

Research at the University of Illinois' Aviation Research Laboratory has begun to be focused on the effects of stress on various heuristics and biases. (See, for example, Wickens, 1987; Wickens & Flach, 1988; Wickens, Stokes, Barnett, & Hyman, 1988; see also the remarks on Wickens's work in Hammond, 1990a, which refers to Wickens, 1987.) Two recent articles (Raby, Wickens, & Marsh, 1990; Stokes, Belger, & Zhang, 1990) continue this effort. Raby et al. (1990) investigate the role of various cognitive biases introduced from the "heuristics and biases" research program with regard to the effects of work load pressure. They consider such topics as planning/participation, overconfidence in time estimates, cognitive "leveling" and task prioritization, and task tunneling. All of their results are worthy of attention from stress researchers.

For example, Stokes et al. (1990) direct their attention to "Anxiety and Cognitive Strategies in Expert and Novice Aviators." They report the following:

> Results confirmed that appreciable baseline salience bias existed, and demonstrated that noise (an environmental stressor thought to simulate the performance effects of anxiety) increased salience bias, whereas increased workload did not. The results showed clear and stable individual differences in salience bias and also suggested that noise influenced "low bias" individuals proportionately more than "high bias" individuals. In terms of visual attention, subjects were also significantly biassed toward the left hemispace. These results are discussed in the context of resource theory. (Abstract)

A detailed description of the flight simulator used at the Aviation Research Laboratory is provided by Stokes (1991) in *Human Resource Management in Aviation*. In the same volume, Jorna and Visser

discuss the effects of anxiety on performance in a flight simulator. They report that their

> study investigated the occurrence of anxiety, and its effects on the performance of high- and low-anxiety subjects. Instructor ratings were compared with objective measures of flight control. State anxiety was found to be increased, particularly by a differential response of subjects to removal of feedback. Anxiety did not influence deviations from the flightpath, i.e., maintaining heading and altitude, but aileron control proved to be quite different. A surprising result was that instructors significantly favoured the high state-anxious subjects. These higher ratings were not supported by objective measures of performance. (1991, p. 123)

In short, studies of the effect of test-anxiety continue to provide results of dubious value.

Social Judgment Theory

Social Judgment Theory grew out of the theoretical and methodological approach put forward by Brunswik (1943, 1952, 1956; see also Hammond, 1966, in which this conceptual framework is described). Brunswik's ideas were organized within the context of "probabilistic functionalism," embracing uncertainty as characteristic of both the environment and the organism (e.g., Brunswik, 1955). Although originally developed for the area of perception, the Brunswikian framework was adapted by Hammond (1955 et seq.), Brehmer, Stewart, Doherty, and others (see Brehmer & Joyce, 1988, for an overview) to the topic of judgment, including expert judgment (for an early example, see Hammond, 1955). The lens model, in particular, was a primary organizing concept.

A significant development in the approach occurred with the introduction of the lens model equation (LME) by Hursch, Hammond,

and Hursch (1964), thus providing a set of parameters that can be used to ascertains the effects of stress. The lens model equation generally takes the following form:

$$r_a = GR_eR_s + C\sqrt{1-R_e^2}\sqrt{1-R_s^2}$$

where

r_a represents achievement, that is, the correlation between a subject's judgment and the criterion of accuracy for that judgment;

G represents achievement when the linearly predictable uncertainty in the environmental system and the subject's response system is removed;

R_e equals the linearly predictable variance in the environment; and

R_s equals the linearly predictable variance in the subject;

C equals the correlation between the residual variance in the environmental system and the subject's prediction system, and thus represents the correlation between nonlinear components of each system.

The general framework of SJT, the lens model, and the lens model equation have been applied to a variety of problems, including multiple cue probability learning (MCPL), interpersonal learning, conflict resolution, forecasting, child abuse, and the effects of psychoactive drugs, among others (see Brehmer & Joyce, 1988, for a review).

Significance for the study of stress. The parameters of the LME can readily be tested for sensitivity to stress, or more precisely, to

change in endogenous conditions. Rothstein (1986), for example, examined the effects of time pressure and found that achievement (r_a) decreased in a MCPL task. He was able to use the LME, however, to discover that the decrease in performance was due to a decreased R_s (consistency in the subject's execution of his/her judgments) rather than to a disorganization in the the subject's knowledge (*G*). That is, the subject continued to do the right thing; it was the consistency of execution that changed. If Rothstein's results are correct, they carry a certain significance, for they show (a) the utility of the LME for discovering the impact of disruption on the judgment system of subjects in terms of *knowledge* (*G*) and control (R_s) (see Hammond & Summers, 1972). This result parallels (but is different from) that obtained by the use of Signal Detection Theory in which accuracy is separated from the decision criterion.

In addition to separation of a change in execution (R_s) from a change in knowledge (*G*) in performance (r_a), the LME also suggests to the investigator the possibility that the uncertainty in the task itself (R_e) may be changed. If objective uncertainty has increased, then (decreasing R_e) the upper limit of performance will be reduced, inasmuch as r_a cannot exceed R_e. This limit is of considerable importance in the evaluation of performance (r_a). For when the task conditions are changed by disruption, performance is apt to be evaluated in terms of perfect performance, when, in fact, the upper limit of performance may be much lower. Without taking the value of R_e into consideration, a faulty conclusion would follow.

Information Integration Theory

Information Integration Theory was developed by Norman Anderson in the early 1960s (for an overview, see Anderson, 1979, 1981, 1982; for recent developments, see Massaro & Friedman, 1990). The approach grew directly out of early work in psychophysics which investigated how organisms perceive stimulus information and sought to identify simple mathematical rules to describe the relationship between stimuli and perceived sensations. For example, the physicist Gustav Fechner (1860/1966) was the first to propose a lawful relationship between a measurement of a physical stimulus (say, the loudness of a sound) and a measurement of an inner mental experience (the perceived loudness of the sound). The relationship (known as Fechner's Law, now replaced by Stevens' Power Law) takes the form $S = k \log I$, where S is sensation, k is a constant, and I is the physical intensity of the stimulus, and it can explain why ever larger outputs in stimulus intensity are required to obtain corresponding changes in perception.

Information Integration Theory attempts to do precisely what Fechner did in psychophysics, but for a broader range of more complex tasks in cognition and social perception: examine how people subjectively evaluate the stimuli that impinge upon them and uncover the mathematical rule or "cognitive algebra" that describes how these perceptions are integrated into a response or judgment. For example, when studying impression formation, one can ask under what conditions subjective ratings on different attribute dimensions are added together or averaged together to form an overall impression of a person (Anderson, 1965). In order to do this,

Integration Integration Theory places considerable emphasis on measuring, in a precise metric sense, the social as well as physical judgments of people, and therefore also owes an intellectual debt to such pioneers in psychological measurement techniques as Thurstone (1931).

Experiments in the Information Integration Theory tradition rely on the concept of functional measurement, which presupposes that psychological values are measurable and can be represented numerically. A typical experiment focuses on measuring the psychological values of the stimuli and the psychological value of the response or integrated judgment. By having subjects evaluate several targets with stimulus values that are manipulated in a factorial design, the pattern of responses can be compared to the pattern which would be expected from various algebraic rules such as addition, averaging, or multiplying. Anderson and his colleagues have used this paradigm to demonstrate that simple algebraic models can describe information use in a wide variety of complex judgments. For example, Anderson (1962) showed that impressions of likableness could be described as the sum of the subjective values for two adjectives that described a person. Shanteau (1974), had people judge the worth of lottery tickets with different probabilities and values, and was able to show that their judgments followed the multiplicative law of subjective expected utility: subjective expected utility = subjective value x subjective probability. And, Lopes (1976), in a study of betting behavior in five-card stud, found that subjects' subjective probabilities of winning followed a conjunctive multiplication rule: the subjective

probability of beating two opponents, A and B, was well-described by the formula subjective probability of beating A x subjective probability of beating B. A potential drawback of these models, however, is that, regardless of how closely the judgments and algebraic models match, the actual form of the underlying cognitive mechanisms may be quite different from the algebraic model (for an example see Graesser & Anderson, 1974).

Significance for the study of stress. Since Information Integration Theory researchers are in the business of measuring subjective values and judgments and testing algebraic relations between them, they provide a wealth of parameters and process components in a variety of domains which might be affected by stress. A unique advantage of the approach is that it can provide baselines against which the effects of stress can be measured even when there is no established normative model. For example, there is no accepted "normative" procedure for forming an impression of another person; yet, judgments can be measured against a well-specified additive model.

Information Integration Theory predicts that stress can have two effects on a given judgment task. First, stress could change the subjective value or judgment attached to different pieces of incoming stimulus information. For example, consider the Lopes study of betting behavior: stress could affect the subjective value of beating any particular opponent. Second, stress could affect the integration rule by which subjective stimuli are combined: under stress, the subjective probability of winning in the Lopes study may no longer closely follow a

multiplicative function of the individual probabilities, but might diverge from the multiplicative rule do to errors or even follow an entirely different algebraic rule. Moreover, these two effects are independent: it is entirely possible, for example, that stress may change all of the subjective stimulus values but have no effect whatsoever on the algebraic relationship between them.

Informal Models

Payne, Bettman, and Johnson (1988)

John Payne has been a pioneer in the development of methodology and theory in the study of multiattribute choice for both risky and riskless options (Payne, 1976, 1980; Payne & Braunstein, 1978). In a typical multiattribute choice study, subjects are asked to choose between several alternatives (say, different apartments for rent) which vary on several attributes (rent, distance from work, etc.). Information about attributes is concealed behind the windows of an array, and subjects reveal the information they are interested in as they go about making their choice. By observing the order in which information is revealed and collecting think-aloud protocols, the cognitive process or processes underlying the choice can be determined. Subjects may, for example, search all of the attributes for each item before going on to the next item, instead search all of the items on a single attribute and then go on to a second attribute, or adopt any of a number of simplifying strategies such as elimination by aspects (Tversky, 1972), satisficing (Simon, 1955), equal weighting

(which ignores probabilities for risky alternatives), or a lexicographic rule which focuses on comparing only the most important attributes.

The observation that individuals may employ a variety of different strategies for any given choice problem (Payne, 1982; see also Abelson & Levi, 1985) led Payne, Bettman, and Johnson (1988) to develop and test an adaptive theory of strategy selection for risky choice in which problem complexity and time pressure (both of which may be considered to be stressors) play a large role in determining which strategy will be selected from the cognitive toolbox. In their model a decision maker attends to both the expected costs (primarily the mental effort required to implement the strategy) and expected benefits (primarily the ability of the strategy to make the best choice) of available strategies during strategy selection. In any given decision environment it is proposed that people will choose the strategy which maximizes accuracy while minimizing effort. Therefore, if people have available multiple strategies which are more or less equally accurate, they will adopt the one which requires the least effort (i.e., the one which is most efficient). The model therefore, unlike heuristics and biases theories, provides a method of predicting which heuristics are most likely to be employed in a given situation (although prediction of exactly which heuristic is most likely to be used requires knowledge about the decision maker's relative values for accuracy and effort). It is also proposed that the accuracy of different strategies varies widely depending on such task characteristics as complexity and time pressure, and that people are adept at recognizing these differences

and at switching from one strategy to another as task conditions warrant.

To test their theory, Payne, Bettman, and Johnson (1988) ran a computer simulation which compared the relative accuracy of 10 choice strategies (ranging from compensatory strategies which consider all of the information to simplified noncompensatory strategies such as elimination by aspects and lexicographic strategies) as complexity (the number of alternatives and atttributes) and time pressure were varied. The simulation was run on a variety of risky choice problems in which subjects chose between gambles with multiple possible outcomes which varied in probability. The simulation demonstrated that, under time and complexity constraints, several heuristic strategies were more accurate than a truncated normative strategy (which tries to integrate all of the information but runs out of time). In addition, the accuracy of any given heuristic was shown to depend on the structure of the decision problem (e.g., the dispersion of probabilities across attributes or the presence of dominated alternatives). No heuristic worked well for the entire space of potential choice problems, demonstrating that multiple strategies are necessary to achieve accuracy in many different situations and contexts. A companion study of human decision makers found that people made choices consistent with the patterns of efficient processing identified by the simulation.

Significance for the study of stress. One implication of the Payne, Bettman, and Johnson approach for the study of decision making under stress is that their work suggests a new baseline for

examining accuracy. The normative criterion is no longer accuracy at any cost but "adaptiveness." Choice behavior under time and complexity constraints should not be compared to the full rational model but to a truncated one. Normative baselines are good criteria for accuracy only when time and mental resources are unlimited and when tasks are simple. People may therefore perform less well under stressful circumstances than under ideal ones but still be performing as well as they possibly can, given the environmental constraints.

The Payne, Bettman, and Johnson framework also provides an explanation for perceptual narrowing and filtration of information, which has often been suggested to be the most typical response by decision makers to stress (see Yates, 1990). People examine less information because it is adaptive to do so. Those strategies which process all of the choices on only a limited number of attributes (e.g., the elimination by aspects or lexicographic strategies) yield the most accurate judgments under time pressure. For choice problems, staying with the normative model (which requires processing all of the attributes for all of the choices) as time pressure increases would not be the optimal strategy.

A close examination of the time pressure results of Payne, Bettman, and Johnson indicates that there may be a "heirarchy" of responses to stress. Under slight time pressure their subjects stayed with the normative compensatory strategy but tried to process information faster. As time pressure increased to moderate levels, they began to narrow their focus, looking at only part of the available information. Finally, under severe time pressure they completely

abandoned the compensatory strategy and switched to heuristic short cuts.

The Payne, Bettman, and Johnson approach is similar to the heuristics and biases approach in that it assumes that coherence is usually poor but that empirical accuracy is often quite good anyway. Under time pressure coherence is lost, but it should be lost (and in fact must be lost) to maintain accuracy. Overall, their model is one of a decision maker who is both coherent and accurate given unlimited time and mental resources and simple problems and who adapts quite well to stress by filtering information and changing strategies. Although accuracy decreases relative to a full normative model, it is maintained relative to the best that can be done under the circumstances. However, it should be noted that they have thus far studied only a single judgment task, multiattribute choice, which people perform very well on when not stressed. It is not clear how their framework would apply to tasks which people typically perform poorly (or even less than optimally) under the best of circumstances.

Einhorn and Hogarth (1987); Hogarth and Einhorn (1989)

Hillel Einhorn and Robin Hogarth recently put forward a model of belief updating which may prove to be extremely valuable in guiding empirical studies of the effects of stress on the cognitive processes involved in evaluative judgments (Einhorn & Hogarth, 1987; Hogarth & Einhorn, 1989). The Einhorn-Hogarth model describes how new information (e.g., an update on the position and course of an unidentified aircraft) is combined with prior beliefs or expectations

about the state of the world (say, whether the aircraft is judged a friend or a foe) to form a revised opinion.

The model assumes that a judge's final revised opinion will depend on five characteristics of the particular task they are facing: (a) task complexity, (b) the amount of information presented, (c) whether the new information confirms or contradicts prior beliefs, (d) the order in which the information is received, and (e) whether the response mode encourages beliefs to be revised in a step-by-step fashion (beliefs are updated after each new piece of information is encountered) or an end-of sequence fashion (beliefs are updated only after all the new information is encountered). Furthermore, the cognitive process employed to arrive at a revised opinion is hypothesized to depend on task characteristics as well. For example, when the response mode is step-by-step, the model predicts that people will revise their beliefs using the following anchoring and adjustment heuristic:

$$S_k = S_{k-1} + w_k S(x_k),$$

where S_k is the degree of belief in a hypothesis after receiving k pieces of information, S_{k-1} is the prior belief (or anchor), $S(x_k)$ is the subjective evaluation of the kth piece of information, and w_k is the adjustment weight for the kth piece of information. When the response mode is end-of-sequence, however, the model predicts that people will consider all of the new information in aggregate and employ a compensatory strategy to update their beliefs:

$$S_k = S_0 + w_k S(x_1, \ldots, x_k),$$

where S_0 is the initial strength of belief in the hypothesis and $S(x_1, \ldots, x_k)$ is a function, typically a weighted average, of the individual evaluations of pieces of evidence which are received after the anchor. The Einhorn-Hogarth model thus goes a step beyond most heuristics and biases theories by offering a prediction of when the proposed heuristic will and will not be employed. Empirical support for various aspects of the model has been obtained by studies of the expert judgments of auditors (Asare, 1990; Ashton & Ashton, 1988, 1990; Meisser, 1990), Army officers (Adelman, Tolcott, & Bresnick, 1990; Serfaty, Entin, & Tenney, 1989), and Patriot air defense operators (Adelman & Bresnick, 1991).

Significance for the study of stress. The Einhorn-Hogarth belief updating model certainly provides well-defined parameters of cognitive processes which may be affected by stress. For example, stress could induce over- or underweighting of new evidence relative to the weighting employed by an individual in a nonstressful situation. Or, if stress were to change the order in which information is considered, the anchoring and adjustment model might result in very different conclusions. However, the model goes even further by predicting situations in which stress should induce a change from one well-specified cognitive process to a very different but again well-specified cognitive process.

Adelman and Bresnick (1991) studied the latter situation. In their experiments with Patriot air defense simulators, officers were asked to judge the probability that aircraft represented on their computer displays were friends or foes and to decide whether or not

to engage the aircraft. For some of the officers, only one aircraft track appeared on the Patriot display at a time, leaving plenty of time for each new piece of information to be evaluated before the next one appeared. This could be considered to be a "low-stress" situation, although Adelman and Bresnick conceptualized their study as manipulating response mode rather than stress. For the other officers, between 5 and 10 aircraft tracks appeared on the display at once, so there was not sufficient time to monitor each track continuously. This is analogous to a "high-stress" situation, since task complexity and time pressure are increased. Adelman and Bresnick predicted that the single-track (low stress) condition would encourage a step-by-step response mode, resulting in an anchoring and adjustment process, whereas the multiple-track (high stress) condition would encourage an end-of-sequence response mode, resulting in a weighted average updating process. This is precisely what happened: under low-stress conditions the Patriot officers exhibited a sizeable anchoring bias but this bias disappeared under high-stress conditions! As this example shows, the Einhorn-Hogarth model makes an interesting and counterintuitive prediction for belief updating under stress, namely, that time pressure during information acquisition will lead to more global, compensatory processing later, resulting in the elimination of an anchoring bias. Judgments therefore become more accurate under stress because it becomes impossible to employ a preferred heuristic strategy which is biased.

It should be noted that, although this prediction is exactly the opposite of the prediction that would be made by Payne, Bettman, and

Johnson (1988), the situation here is quite different. In the Adelman and Bresnick study, there was time pressure during information acquisition but when it came time to make a decision there was sufficient time to review all of the information that had been received and apply a compensatory model. (See also the remarks above on Signal Detection Theory; Harvey et al., in press, also observed performance to improve under stress.)

Cognitive Engineering

Cognitive engineers such as Vicente and Rasmussen, as well as systems analysts, have largely ignored the "gloomy conclusions" reached by psychologists; they believe, as do most economists, that these conclusions have little or no meaning outside the psychologists' laboratory. But cognitive engineers share with human factors researchers an interest in "human error" (see Reason, Moray, and Hancock above). For example, Vicente and Rasmussen (in press) state that "the single most important concern in improving system safety is to provide operators with the support required to deal with unfamiliar and unanticipated abnormal situations." That is not the conclusion that would be drawn by a reader of the current J/DM and social psychological literature. That reader would conclude that it is the *familiar* and the *anticipated*, the everyday, normal situation that evokes errors in judgments. Indeed, that is the situation most frequently studied by heuristics and biases researchers. Can both conclusions be true? No doubt both *are* true; the question to be addressed by both the design engineers who must prepare for the appearance of the unfamiliar, and the psychologists who study judgments of the familiar

is: What are the common features of the everyday, anticipated situations *and* unfamiliar, unanticipated situations that lead to fallacious *and* accurate judgments? Currently, the research treatises that study these circumstances in isolation do not lead us far because they do not assist in answering the fundamental question: What circumstances induce which forms of cognitive activity? And which circumstances produce the cognitive activity desired? Which task circumstances are most disrupting? With what cognitive consequences?

What is needed, therefore, is a theory that will enable us to assist the design engineers to provide support for operators when they are "required to deal with unfamiliar and unanticipated situations," as well as to assist those who wish to assist those persons who are "operating" in familiar, well-anticipated situations. And it is that theory that will be required if we are to ascertain the effects of various stressful conditions on cognitive activity.

Note that the cognitive activity desired may not be the same in both sets of circumstances. Generally, if not always, the "operators" that the design engineers Vicente and Rasmussen have in mind are to be supported in their efforts to engage in rational, defensible cognitive activity. Thus analytical, possibly rule-bound, cognition will be the goal—*if* that is attainable. If it is not, then some *other* form of rationality will have to be specified and defended—perhaps in a court of law as well as in scientific dispute. For example, suppose one argues that it is not possible to employ fully analytical, rule-bound behavior under certain task circumstances; what then?

It is precisely this situation that makes clear the need for theories of judgment that focus neither on purely analytical nor purely intuitive forms of judgment and decision making. For in many engineering situations that create unfamiliar unanticipated conditions time pressure will not permit analysis, and responsibility will not permit pure intuition. What is needed are theories that focus on cognitive activity that contains elements of both.

Theories of this type have been present for some time. The best known is Simon's theory of bounded rationality, and the Brunswik/Hammond theory of quasirationality. Not to be overlooked is the judiciary's use of the term "reasonable," as used, for example, in the expression, "Was the defendant's conduct (or judgment) reasonable in light of the circumstances?," a question often raised in connection with judgments made under stressful conditions, as might occur in military operations, the operation of a vehicle, or airplane, or ship, etc. (Consider, for example, the captain of the Exxon Valdez absenting himself from the bridge prior to the ship running aground while exiting Prince William Sound; was that "reasonable"? The jury thought so.)

The central question is: If retraceable, logically, empirically defensible judgments *cannot* be carried out, what criterion should be used to evaluate the judgment that was applied? The court's "reasonableness" standard is simply too vague, too situation-bound (juries regularly disagree) and too often found (later) to be determined by irrelevancies. *Should* the above theories be applied? *Can* they be applied? *How* can they be applied? And in view of Ballou and Pazer's

(1990) observations that some decisions are more satisficing than others, is it possible to argue that one "satisficing" action was more satisfactory than another? Or that one was later found to be "unsatisfactory" although "satisficing" at the time?

We now turn to a consideration of each of the theories that directly address cognitive activity that falls under the rubric of bounded rationality and/or quasirationality.

Bounded Rationality

What is to be learned from Simon's concept of bounded rationality in relation to stress?

First, recall that bounded rationality is not irrationality; it is rationality within bounds, that is, within a problem space that is less than complete, less than that demanded by unbounded rationality. Most important the problem space is restricted or diminished under *ordinary*, that is, nonstressful, conditions by virtue of human limitations on information processing as well as by the level of aspiration. Stress, then, might (a) increase the problem space, (b) decrease it, or (c) leave it unchanged. Theory is no guide here, nor can we find studies that offer empirical evidence of the possible effects of stress on bounded rationality. And it is easy to think of circumstances that might produce any of the three results indicated.

Simon directly addresses the question of the effects of stress on cognition in "Motivational and Emotional Contexts of Cognition" in *Models of Thought* (1979) in which he notes that "if real-time needs are to be met, then provision must be made for an *interrupt system.*" An interrupt system must have two properties: (a) "a *noticing*

program" which will have the capacity "to *notice* when conditions have arisen that require ongoing programs to be interrupted"; and (b) "the noticing program must be capable of *interrupting* and setting aside ongoing programs." Simon also notes that drive level will have to be considered, as well as "a threshold for each drive [that] determines at what drive level the goal becomes 'urgent' and interrupts the ongoing program" (p. 34). Although Simon acknowledges that interruption is not the only circumstance that evolves emotion, interruption is most germane to the types of circumstances that involve J/DM and/or problem solving. Because interruption is a common experience and, for the most part, is unwelcome in situations demanding cognitive activity—particularly when time pressure is present—this suggestion is relevant to our purposes. The next question is what are the cognitive consequences of interruption? Simon observes and quotes Hebb (1949, 238-240; 250-258) to the effect that "the emotional stimulus is to be regarded as more often *interrupting* than *disrupting* behavior" (p. 35). Moreover, "the responses to interruption are largely *adaptive*" (italics added). Thus "interruption . . . may evoke an elaborate goal-oriented chain of activity (e.g., the reactions of a trained soldier to the sound of approaching aircraft)" (p. 35). But "when the emotion-producing stimuli are persistent as well as intense, they sometimes become disruptive and produce nonadaptive behavior" (p. 35). Simon also suggests that people sometimes do and sometimes do not learn to cope with emotion-producing stimuli (p. 36).

These remarks are somewhat disappointing: The question is: What forms of interruptions are likely to produce which of these results? And how would these results compare with those produced by unbounded cognitive systems? What difference will interruptions make for large scale cognitive systems vs. small scale for the well-trained expert vs. the novice? And what difference does it make whether interruptions occur early or late in problem-solving work?

Simon's brief treatment of emotion does not take us far with respect to anticipating the effects of stress on bounded rationality. (But see Simon, 1983, on the role of emotion in creativity.)

Summary

The formal models (e.g., SDT or SJT) described in this section offer parameters that might well turn out to be sensitive to the effects of stress. For example, (a) Harvey, Hammond, Lusk, and Mross (in press) and Lusk (in press) have demonstrated the sensitivity of SDT parameters to stress in relation to weather forecasters' predictions, and (b) Rothstein (1986) and Schwartz and Howell (1985) have shown the effects of time pressure on inducing different forms of cognitive activity as predicted by CCT. Stokes and Raby (1989) have demonstrated that the use of certain heuristics is affected by stress in aviators; Payne et al. (1988) have shown that the use of certain strategies is affected by time pressure; Hogarth et al. (1991) have found subjects' cognitive activity to be influenced by the "exactingness" of the task where there are heavy penalties for mistakes. In short, there are some empirical results that give us

reason to believe that J/DM models will be useful in the study of the effects of stress. But these results are certainly skimpy, and none have been replicated. In short, the use of J/DM models in stress research rests on plausibility arguments—so far.

Is it plausible to urge that these models—despite the limited evidence of their usefulness—be employed in stress research? In our view, there is no other choice. At present, the great majority of research that examines the effect of stress on cognitive activity employs no models of cognitive process at all. If process is to be understood—and, in our view, it is essential that it be understood—models of J/DM research programs will have to be introduced. It is to be hoped that if introduced in a *systematic* manner, they will considerably advance our understanding of this topic.

5. Methodology

This section is devoted to methodology—the science of method—rather than to a discussion of specific methods. Discussion of methodology thus entails the examination of a body of methods—such as idiographic vs. nomothetic methods—rather than the examination of a particular method such as analysis of variance, multiple regression, etc. Although the distinction has been almost entirely ignored (due to the debasement of the concept of methodology, often to mean technique) it is important to maintain it here in order that we differentiate among the *sets* of methods employed to pursue the study of J/DM.

The primary methodological question that has emerged over the past decade—now that the previous three or four decades have made it obvious that J/DM *can* be studied—is whether the methods conventionally employed in J/DM studies will permit the generalization of results beyond the specific, usually laboratory, conditions within which the results have been obtained. In short, generalization has become a critical methodological issue, to which we turn first.

Generalization and Representativeness of Task Circumstances

The prominence of the matter of generalization can be seen in the program statement for the Decision, Risk, and Management Science Program within the National Science Foundation:

> Research supported by the DRMS Program should have relevance to *operational context*, be grounded in *theory*, be based on *empirical observation* or be subject to *empirical validation*, be *generalizable*, and incorporate *social, behavioral*, or

> *organizational aspects* of operational processes and decision making.

Generalization of results is thus stated as an important criterion for support of J/DM research by NSF. (See Hammond, 1986a, for a commentary on this criterion.)

The reason for the concern expressed in the NSF program statement lies in the fact that doubts about the generality over tasks have been expressed for some time. These doubts are based on the degree to which the laboratory task circumstances, or task properties, are representative of the task conditions toward which the generalization is intended. First raised in persistent fashion by Brunswik (1943), these doubts persist. For example, Smith and Kida (1991) recently reviewed some 25 studies of accountants' J/DM behavior and concluded that "the results of audit-judgment research indicate instances in which biases (see heuristics and biases approach above) found readily in other research are not evident in the judgments of professional auditors" (p. 485). They further note that, "important issues concerning the generalizability of these findings [concerning biases in judgments] have been raised" by several researchers. Smith and Kida (1991) thus challenge the generalizations about heuristics and biases on the ground that they are "based on studies that, in large part, use tasks and subjects that are not representative of the contexts and individuals for which these findings seem most relevant" (p. 485).

On the basis of their broad-ranging review of studies of expertise Ericsson and Smith (1991) make a parallel observation about the lack of representativeness of tasks used by researchers:

> In most other complex task domains, such as physics and medical diagnosis, investigators tend to select a small number of tasks without specifying the population from which those tasks were chosen to be a representative sample. One reason for this is that a detailed task analysis of even a single complex problem is difficult and extraordinarily time-consuming." (p. 15)

Criticism of tasks used by psychologists frequently becomes extended into questions about the restricted nature of laboratory research itself; unfavorable comparisons between the "impoverished" conditions of laboratory tasks and the complexity of "real life" soon follow. This point is frequently made by researchers in the field of ergonomics. Moray, for example, has been especially outspoken, and because he identifies himself as a former experimental psychologist his remarks are particularly significant:

> There has been a great deal of interest since the early 1950s about the nature of attention (Broadbent 1958, Moray 1969, Kahneman 1973, Wickens 1984, among others). But the majority of this work has been performed in laboratories under circumstances which are only marginally relevant to the limits of attention in real tasks. (1990, p. 1203)

In discussing research on transport systems he states:

> Given such an impoverished sampling [from laboratory studies] of the environment, and looking at the rates at which people operate transport systems, how do they manage at all? The answer is three-fold. In the first place, the environment is highly structured. A glance at one part of the world tells us not merely what its state is, but also a great deal about neighbouring regions. The redundant structure of the world allows correlation to substitute for observation. Secondly, the bandwidth of the world is limited. In the case of transportation systems in particular, the bandwidth of the world is generated by the rate at which a vehicle traverses the environment. The faster I drive, the greater the rate of change of the visual array, and the greater the bandwidth of information generation. But it always remains low compared with the bandwidth of the visual system. Thirdly, the paths taken by vehicles are severely constrained by physics. The paths may be constrained to one dimension, as in a train, to two dimensions for surface-effect vehicles, or to three dimensions for submarines or aircraft. The degrees of freedom of motion may reach as high as 6 degrees for

> rotorcraft. But physics constrains motion to rather predictable paths for most of the time. Hence the motion of vehicles is to a great extent predictable. (p. 1204).

He concludes:

> It seems that we are in a position to model the impact of mechanisms of data acquisition in human operators rather well, although few examples of complete modelling have been attempted. This is largely because of the impoverished laboratory environments used in research on detection, perception, attention, and workload. As we increase the realism of laboratory studies, and move toward field studies with improved instrumentation for eye movements and ways of measuring operator behaviour other than direct performance, we will find, I believe, that we have quite a strong armamentarium for attacking these problems and predicting the modes and causes of human error. (p. 1211)

Similar attitudes toward the restrictive nature of laboratory research can be found in articles by design engineers such as Rasmussen (1986) and others.

The problem of generalization from laboratory studies has also been discussed in the context of "field vs. laboratory research," the most recent example being the contents of the book by Klein, Orasanu, and Calderwood (in press). In their chapter, Orasanu and Connolly (in press) list eight characteristics of "naturalistic" decision making that they assert "frequently are ignored in decision making research":

1. Ill-structured problems
2. Uncertain dynamic environments
3. Shifting, ill-defined or competing goals
4. Action/feedback loops
5. Time stress
6. High stakes
7. Multiple players
8. Organizational goals and norms

They further assert that

> phenomena observed in complex natural environments may differ substantially from those observed in the laboratory based on decontextualized tasks performed by novices with little stake

> in the outcomes. Ideally, laboratory tasks involve micro environments that abstract essential variables from the broader environment to which the experimenter wants to generalize (See Hammond [et al.], 1980, and this volume for a discussion of Brunswik's representative design). In fact, critical variables may be missing or changed in the lab, with major consequences for the behavior of interest.

Orasanu and Connolly also claim that "naturalistic decision making research has yielded new findings:"

> In naturalistic dynamic settings, experts frequently generate and evaluate a single option rather than analyze multiple options concurrently.
>
> Experts are distinguished from novices mainly by their situation assessment abilities, not their general reasoning skills.
>
> Because most naturalistic decision problems are ill-structured, decision makers choose an option that is good enough, though not necessarily the best.
>
> Reasoning is "schema driven," that is, guided by the decision maker's knowledge, to search and assess information, and to build causal models of events.
>
> Deciding and acting are interleaved.

Criticism of laboratory research can also be found in the field of memory research. Neisser (1978) observed "that the naturalistic study of memory is an idea whose time has come" (p. 3). Other well-known psychologists, including Barker (1968), Bartlett (1932), Bronfenbrenner (1974), Brunswik (1956), Chi, Glaser, and Farr (1988), Gibson (1979), and Lorenz (1966) among others have expressed heavy criticism of laboratory methods. Brunswik, in particular, made detailed criticism of laboratory research a major part of his efforts (see, for example, *Perception and the Representative Design of Psychological Experiments*, 1956).

Arguments for the representative design of experiments are seldom, if ever, answered with the seriousness and detail with which they are put forward, which may account for the growing frequency with which Brunswik's argument is repeated without citation or even

with an incorrect citation (see, for example, Smith and Kida, 1991, p. 472, who attribute the representative design argument to Ward Edwards). But Gibson and Brunswik, and their arguments favoring ecological psychology in contrast to laboratory research, are now receiving increased attention, particularly from those in the field of engineering and design (see Vicente, Rasmussen).

Kleinmuntz (1987) was careful in his treatment of the problem of generalization in his "bridging" chapter in Hancock's *Human Factors Psychology* (1987).

> One conclusion sometimes drawn from this body of research is that people are generally inept decision makers, whose feeble cognitive abilities fall far short of what they believe them to be. . . . Some caution must be exercised before accepting this inference. For instance, some researchers have questioned the external validity of laboratory demonstrations of biased decisions. While Slovic et al. (1977, pp. 14-17) reviewed a number of studies that showed similar findings in laboratory and field settings, Ebbesen and Konecni (1980) described a number of instances where laboratory research failed to generalize. Einhorn and Hogarth (1981) suggest:
>
> > "The issue of external validity is not liable to be resolved without recourse to theory that attempts to answer how tasks vary between the laboratory and natural environment and what kinds of effects can be expected from such differences." (p. 81)
>
> This theoretical framework would permit an assessment of the *general* conditions under which heuristics work well or badly as well as the distribution of those conditions in the natural environment.
>
> This framework may also guide attempts to "debias" judgment. Fischhoff (1982) suggests a classification scheme for efforts to improve or correct judgment processes. One category of efforts aims to correct problems caused by the task, by either removing unfair task demands or clarifying misunderstood elements of the task. Another category finds fault with the judge's limited cognitive capacity, with the recourse being either to engage in training exercises or to replace the judge with a mechanical aid. A final category involves mismatches

between the task and the judge; the judge has the requisite cognitive abilities to perform the task, but for some reason those abilities are not used. Possible corrective strategies include restructuring the task through decomposition or alternate formulations.

A final corrective option worth considering is formal training in statistics. Nisbett, Krantz, Jepson, and Kunda (1983) demonstrate that some judges can and do use *statistical heuristics*, simple judgment rules that are roughly equivalent to formal statistical principles. Furthermore, they show that training in statistics increases both the use and the effectiveness of these rules. Further investigations of the origin and extent of these statistical heuristics promise to provide important insights into the potential for avoiding biased judgments.

Given the current state of knowledge about heuristics, there is no alternative to an intuitive assessment of the prevalence of bias. Christensen-Szalanski and Beach (1984) suggest that those who conclude that severely biased judgments are widespread are victims of a "citation bias," a propensity to pay more attention to the published evidence of poor performance than the evidence showing good performance. They go on to suggest that this bias is a passing fad in the literature. While the notion of judgment researchers falling prey to one of their own biases is certainly ironic, it is not at all clear that focusing attention on systematic errors is a mere fad. For instance, Evans (1984) points out that instances of poor judgment are important since when heuristics do fail, judges are usually not aware of the failure (also see Einhorn & Hogarth, 1978; Kahneman & Tversky, 1973). Furthermore, focusing on errors is an explicit research strategy that has proven useful in other domains of cognitive psychology. As Tversky and Kahneman (1983) point out, studies of "cognitive illusions" are diagnostic of basic judgment processes in the same way that studies of visual illusions are diagnostic of basic perceptual processes.

> "The focus on bias and illusion is a research strategy that exploits human error, although it neither assumes nor entails that people are perceptually or cognitive inept. In cognition, as in perception, the same mechanisms produce both valid and invalid judgments." (p. 313). (Kleinmuntz, 1987, pp. 125-127)

Despite such disclaimers, Gigerenzer (in, e.g., Gigerenzer et al., 1991) has been very critical of the generalizations drawn by heuristics and biases researchers because of the task materials used by them:

In several Bayesian-type studies of revision of belief, representative (random) sampling from a reference class is a crucial issue. For instance, Gigerenzer, Hell, and Blank (1988) showed that subjects' neglect of base rates in Kahneman and Tversky's (1973) engineer-lawyer problem disappeared if subjects could randomly draw the descriptions from an urn. Similar results showing people's sensitivity to the issue of representative versus selected sampling have been reported by Cosmides and Tooby (1990, August), Ginossar and Trope (1987), Grether (1980), Hansen and Donoghue (1977), and Wells and Harvey (1977), but see Nisbett and Borgida (1975).

This study has also demonstrated that judgments of single events can systematically differ from judgments of relative frequencies. Similar differences were found for other kinds of probabilistic reasoning (Gigerenzer, 1991a, 1991b). For instance, the "conjunction fallacy" has been established by asking subjects the probabilities of single events, such as whether "Linda" is more likely to be (a) a bank teller or (b) a bank teller and active in the feminist movement. Most subjects chose the latter, because the description of Linda was constructed to be representative of an active feminist. This judgment was called a conjunction fallacy because the probability of a conjunction (bank teller and feminist) is never larger than the probability of one of its constituents. As in the engineer-lawyer problem, the representativeness heuristic was proposed to explain the "fallacy." Fiedler (1988) and Tversky and Kahneman (1983), however, showed that the conjunction fallacy largely disappeared if people were asked for frequencies (e.g., "There are 100 persons like Linda. How many of them are. . . ?") rather than probabilities of single events. Cosmides and Tooby (1990, August) showed a similar striking difference for people's reasoning in a medical probability revision problem. The subjects' task was to estimate the probability that people have a disease, given a positive test result, the base rate of the disease, the false-alarm rate, and the hit rate of the test. Originally, Casscells, Schoenberger, and Grayboys (1978) reported only 18% Bayesian answers when Harvard medical students and staff were asked for a single-event probability (What is the probability that a person found to have a positive result actually has the disease?). When Cosmides and and Tooby changed the task into a frequency task (How many people who test positive will actually have the disease?), 76% of subjects responded with the Bayesian answer. These results suggest that the mental models subjects construe to solve these reasoning problems were highly responsive to information crucial for probability and statistics—random versus selected sampling and single events versus

> frequencies in the long run. (Gigerenzer et al., 1991, pp. 524-525)

Thus once more it is evident that generalizations about the quality of J/DM are being sharply challenged because of the nature of the experiments that are conducted. Brehmer and Dörner (1991) state the issue clearly:

> Psychology lives with many tensions. One is that between research in the laboratory and research in the field. Field researchers criticize laboratory research for lack of relevance or "ecological validity" as it is now often called. Laboratory researchers, on the other hand, criticize field researchers for lack of control with attendant problems in making causal interpretations of their findings. Both are right, of course, especially in the eyes of the applied psychologists who try [to] use the knowledge produced by psychological research.
>
> The root of these problems lies in the inability to handle *complexity*. In field research, there is often too much of it to allow for any more *definite* conclusions, and in laboratory research, there is usually too little of it to allow for any *interesting* conclusions.
>
> Perhaps the most important difference between field research and laboratory research lies in the very *different picture of man* that emerges from the two forms of research. Laboratory research yields a picture of man as rigidly controlled from the outside. Field research on the other hand, and this is especially true of field research in the French tradition (see, e.g., de Keyser, Decortis, & van Daele, 1988) gives a picture of man as a selfregulating system. This difference is not surprising, and it simply reflects the kind of person who can be observed in the laboratory and the field. Again, this difference is best understood in terms of differences in complexity; laboratory experiments are set up to produce clear cut relations between independent and dependent variables, and they therefore cannot use experimental situations that are complex enough to allow subjects to exhibit the kind of self regulation observed in field studies. *This gap must be bridged; psychology cannot, in the long run, live with two such different conceptions of man* [italics added]. (pp. 1-2)

In an effort to represent task circumstances faced by meteorologists, Lusk and Hammond (1991) used a sample of weather

situations taken directly from a population of computer displays of weather conditions in order to study expert on-the-job performance. The Brehmer/Dörner approach, on the other hand, is used to construct conditions of interest for theory testing.

It is obvious that research that claims to produce information regarding judgment under stress must stand up to criticism regarding the representativeness, and thus meaningfulness, of experimental conditions for those who wish to make use of the results. If such criticism is to be overcome, the research must claim that the information produced can *somehow* claim generality beyond the laboratory conditions under which it was produced. Stress research will, therefore, in all likelihood reinforce the trend toward representative design.

Idiographic vs. Nomothetic Aims

The growing demand for representativeness in the experiment of the conditions toward which generalizations are intended is linked with a second trend, that of a demand for idiographic rather than nomothetic research goals. Generally speaking, the term idiographic refers to studies of individuals, whereas the term nomothetic refers to law finding over populations. This distinction can be simplified by referring to a demand for the description of the behavior of separate individuals in contrast to the reporting of averages for groups, as in between-groups ANOVA analyses. Fundamentally, the argument for an idiographic analysis is essentially the same as the argument for the representative design of experiments, for both arguments insist that

analyses should be made of the unit (or element or category or sample) toward which the generalization is intended. –

The well-known statistician Yule objected to the all-too-ready use of averages as early as 1921: "Statistical methods . . . should be regarded as ancillary, not essential. They are only essential where the subject of investigation is itself an aggregate, as a swarm of atoms, or a crowd. But here the subject is the individual" (pp. 105-106).

The idiographic/nomothetic distinction was given prominence in personality research by Gordon Allport in 1937 in which he argued for the uniqueness of each personality. About the same time Lewin (1935) objected, specifically and forcefully, to the use of means or averages in the effort to discover lawful relations in psychology, an argument that carries meaning today in relation to the (almost universal) use of significance tests across groups.

The principal problem created by between-subjects nomothetic research is that of generalizing the results to individuals, when, in fact, the results are derived from averages. An interesting example of the difficulty is provided by an epidemiologist (Sackett, 1989) who returns to practicing as a bedside physician. He recognizes the difficulty and challenges himself to defend the application of knowledge derived from between-group studies to the individual patient whom he must treat. He states his problem this way: "First, how should clinicians decide whether the results of a randomized trial . . . apply to their own particular patient? . . . That is, how ought we consider the generalizability or external validity of internally valid randomized trials [experiments]?" (p. 310). It should be of great

interest to students of J/DM that Sackett is unique in facing up to this problem, and, most important, must find his own, logically defensible, solution to it.

Closer to psychology, Lamiell and Trierweiler (1986) meet this issue in a manner that carries direct significance for the methodology of research in J/DM:

> How defensible are traditional aggregate indices of (in)consistency as grounds for generalizations concerning the (in)consistency of individuals? . . . (p. 461)
>
> There is a purely logical difficulty related to the matter of predicting individual behavior with which apologists for the prevailing paradigm [nomothetic] have yet to come to grips. This difficulty resides in the fact that given empirical findings of the sort on which traditional "nomotheticists" are wont to base their assertions about behavioral predictability, one is for all practical purposes never in a position to say anything at all about the predictability of a given individual's behavior.
>
> As regards behavioral predictability, the only knowledge that traditional "nomothetic" inquiry yields is knowledge about the average of the (squared) errors of prediction across individuals. It does not yield knowledge about the error of prediction for *any* one individual, and hence cannot properly be said to yield knowledge about the predictability of *any* one individual's behavior.
>
> In order to claim knowledge about the predictability of an individual's behavior, one must by definition have error-of-prediction knowledge that is interpretable at the level of the individual, and traditional "nomothetic" inquiry meets this requirement only under conditions where the correlation (simple or multiple) on which predictions are based is unity (i.e., +/-1.00). In that and only that event, error-of-prediction is known at the level of the individual precisely because it is by definition zero for each and every individual. When the correlation is less than unity, however, error-of-prediction is by definition indeterminate at the level of the individual and hence unspecifiable for each and every individual. In fact, for any given individual, the error of prediction could take on any value permitted by the scale on which the criterion variable has been defined. No greater precision concerning the predictability of individual behavior can legitimately be claimed given the kind of

> knowledge yielded by traditional "nomothetic" inquiry and this much "precision" can legitimately be claimed before—and whether or not—any "nomothetic" inquiry is ever carried out.
>
> The lesson here, of course, is that if it is the predictability of Smith's behavior that is at issue, then it is knowledge about the predictability of Smith's behavior that one must acquire, and if it is the predictability of Jones' behavior that is at issue, then it is knowledge about the predictability of Jones' behavior that one must acquire. Conventional "nomothetic" inquiry simply does not offer such knowledge, and to note that "traditional wisdom" has long been otherwise is merely to note that "traditional wisdom" has for several decades been rather longer on tradition than wisdom. Moreover, if it now appears to the reader that "nomotheticism's" epistemological inadequacies vis-à-vis the matter of predicting individual behavior stem from the aggregation problem mentioned earlier in our discussion of (in)consistency in personality, that is because they do. . . .
>
> Pared to its essentials, this argument is that while the critical assertions of any theory of personality concern *individual* behavior/psychological functioning, the aggregate statistical findings issuing from studies of individual differences variables have (unless they are perfect, which is never) *no legitimate interpretation of any kind whatsoever at the level of the individual.* (pp. 466-467)

Lamiell and Trierweiler's point (see also Lamiell, 1982) is exactly the same as the one Lewin and others urged upon psychologists a half century ago—with success.

Brunswik made a similar methodological argument in a stronger, more rigorous form; that is, he argued for an idiographic-*statistical* approach, consistent with his plea for a representative design of experiments; each individual should be studied over samples of tasks relevant to the generalization intended. Thus he urged an approach that was essentially the reverse of the conventional one, that of using one task but sampling many individuals. Of course, in the idiographic-statistical approach individuals are sampled, but each subject's behavior is studied individually, that is, idiographically, over a sample

of tasks appropriate to the ecological generalization. Lawful behavior is discovered by ascertaining how many and which individuals exhibit the lawful behavior in question over the ecology included, or sampled, in the study, a methodological position taken by Norman Anderson as long ago as 1962.

Meehl (1990) has recently supported Brunswik's assertion that psychologists are careful to sample subjects to support generalizations but fail to apply the same logic to situations: "One badly neglected sampling problem, probably more important than the sampling of organisms to which such meticulous attention is conventionally paid, is the sampling of situations, which should be in some sense 'representative' of the statistical ecology of the species studied" (p. 41).

Idiographic research aims at discovering the behavior of single individuals over a series of tasks—sometimes formally sampled, sometimes not—and thus *generalization is intended to be achieved over specified task conditions*. Idiographic-*statistical* research goes further; it makes a formal attempt to sample conditions and thus includes *statistical tests* over conditions. (For examples, see Hammond, Hamm, & Grassia, 1986; Hammond et al., 1987; Lusk & Hammond, 1991; Lusk, Stewart, Hammond, & Potts, 1990; Stewart, Middleton, Downton, & Ely, 1984; Stewart, Middleton, & Ely, 1983; and Stewart, Moninger, Grassia, Brady, & Merrem, 1989; for a review see Brehmer & Joyce, 1988.)

Recently, even archaeologists have declared that the idiographic-statistical approach is appropriate for advancing their

discipline. As Mithen (1990) has put it: "To focus on the individual decision maker is to adopt . . . methodological individualism . . . in which explanations for social and economic phenomena must make reference to the dispositions of individuals" (p. 2). Throughout his treatise on the decision making of pre-historic people Mithen makes use of this approach to understand their decision making processes.

Nickerson and McClelland (1991) have shown how the between-subjects (i.e., nomothetic) approach leads to erroneous conclusions when the investigator's intentions are in fact idiographic ones (see also Mandler ref. in Nickerson). Perhaps the most flagrant example of a misdirected generalization due to a between-subject approach is the widespread acceptance of the research on "illusory correlation." Despite the fact that "illusory correlation" has never been demonstrated to occur in any one of the hundreds of subjects studied, the "fact" of illusory correlation has been cited hundreds of times as if it *had* been established (see Hammond, 1978, 1990b for details).

The importance of the distinction has also been noted by the eminent historian of statistical methodology, Ian Hacking (1986), who provides an example of misconceptions generated by averaging—identical in principle to those pointed out by Lewin a half century earlier. Hacking's example is especially useful because it involves a very practical and important engineering problem—that of eliminating "fly ash" from the pollution caused by coal-fired electric-power stations.

> At present, fly ash is eliminated by building a tall chimney with many baffles. The walls and baffles are electrostatically charged, and the fly ash is charged in the opposite direction so that it will stick to the walls and baffles. After a certain amount of ash is

collected, the process is stopped and the walls are banged to dislodge the caked ash, which is rather difficult to remove. This is a simplified version of a procedure in use for more than half a century, designed chiefly not by theoreticians, but by sheet-metal fabricators.

Two quite different phenomena are involved in these chimneys. One is the mechanical, aerodynamic turbulence of the particles going up the chimney, a turbulence greatly increased by the presence of the baffles and the velocity with which the hot gases escape. Precise analysis of this turbulence is difficult, but it belongs to a class of problems long investigated and now well understood in essentials, especially now that computer simulation relieves us of the need to solve all the messy equations that arise. We can construct models, susceptible of calculation and analysis, that describe the passage of particles in turbulent fast flow through a baffled chimney. The second phenomenon is the attraction of the charged particles in motion to the walls and baffles. This phenomenon is also well understood.

We may form the picture of the particles being subject to laws of turbulence and behaving accordingly. We may also form the picture of charged particles being subject to laws of electrostatics and behaving accordingly. The trouble is that the fly ash falls under both laws. The practical assumption, perhaps made implicitly, is that the ash will behave as a sort of average effect of both laws: The ash does its best to obey both laws, satisfied neither master, but to some extent goes along with both. Such an assumption, dressed in less metaphorical but no more precise language, has been the basis for building half a century of coal-fired power plants.

According to my sources, the truth is entirely different. The behavior of fly ash under conditions of turbulence and electric charge is altogether unlike what understanding of either or both sets of laws would lead one to believe. At present, we are at the level of terribly empirical research, trying to observe the behavior of millions of tiny, man-made, uniform particles going along tubes of various shapes and charges. Some tentative models have been proposed. It has begun to look as if the standard power-plant chimney is almost the least efficient way to trap fly ash. It might even have been better to have had no baffles (to diminish mechanical turbulence) or to have had no electrostatic charges, for the benefits of each system interfere with the benefits of the other. More important, altogether different designs could prove to be much cheaper to build, much cheaper to operate, and a good deal better at controlling this kind of pollution. (p. 145)

The discovery of the misuse of averaging in such diverse fields as psychology and engineering is astonishing.

Integrating the Idiographic-Statistical Approach with Representative Design

Campbell and Fiske (1959) long ago demonstrated the fact that psychological research methods were generally often more highly correlated with one another than the traits the methods purported to measure. That is, purportedly *different traits* were found to be highly correlated when measured by the same method (e.g., questionnaires), but low correlations were found among measures of the same trait when they were measured by *different* methods. In short, method variance accounted for more of the subjects' behavior than did trait variance—a result that carried enormous (negative) implications for personality research, but did little to affect the behavior of the researchers.

Hammond et al. (1986) turned the multitrait multimethod (MTMM) approach to the advantage of idiographic-statistical goals and to the representative design of experiments by arranging the conditions of the experiment so that the MTMM analysis could be applied separately to each *individual* (expert) subject in the study. Thus, the authors claim that "a performance validity matrix derived from Campbell and Fiske's methodology provides an analytical, that is, logically defensible method for evaluating generality over concepts [called "traits" by Campbell and Fiske] and methods, as well as criteria of performance" (Hammond et al., 1986, p. 260).

A conventional MTMM matrix is presented in Table 5.1; a MTMM matrix for a single expert is presented in Table 5.2; and a contrast with the customary nomothetic approach is presented in Table 5.3.

Table 5.1
A Synthetic Multitrait-Multimethod Matrix

		Method 1			Method 2			Method 3		
Method	Traits	A_1	B_1	C_1	A_2	B_2	C_2	A_3	B_3	C_3
	A_1	(.89)								
1	B_1	.51	(.89)							
	C_1	.38	.37	(.76)						
	A_2	.57	.22	.09	(.93)					
2	B_2	.22	.57	.10	.68	(.94)				
	C_2	.11	.11	.46	.59	.58	(.84)			
	A_3	.56	.22	.11	.67	.42	.33	(.94)		
3	B_3	.23	.57	.12	.43	.66	.34	.67	(.92)	
	C_3	.11	.11	.45	.34	.32	.58	.58	.60	(.85)

Note: Reliabilities are indicated in parentheses in the main diagonal. Values in lower diagonals (between triangles) are convergent validity coefficients (monotrait-heteromethod) derived by measuring the same trait with different methods. Solid lines delineate heterotrait-monomethod triangles. Dashed lines delineate heterotrait-heteromethod triangles. Values in both triangles are discriminant validity coefficients. From "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix" by D. T. Campbell and D. W. Fiske, 1959, *Psychological Bulletin, 56,* p. 82.

Table 5.2

Coherence Validity Multiconcept-Multimethod Matrix for Artificial Engineer

		Method								
		Filmstrip			Bar Graph			Formula		
Method	Concepts	E	S	C	E	S	C	E	S	C
Filmstrip	E	(.957)								
	S	.035	(.825)							
	C	-.117	.951	(.793)						
Bar Graph	E	(.890)	-.283	-.364	(.996)					
	S	-.244	(.713)	.712	-.491	(.968)				
	C	-.051	.706	(.591)	-.282	.581	(.979)			
Formula	E	(.864)	-.360	-.452	(.985)	-.548	-.337	(.999)		
	S	-.093	(.393)	.427	-.209	(.422)	.452	-.232	(.415)	
	C	.176	.403	(.333)	.015	.270	(.340)	.004	.001	(.315)

Note: Reliabilities are indicated in parentheses in the main diagonal. Values in ellipses (between triangles) are convergent validity coefficients (monotrait-heteromethod) derived by measuring the same trait with different methods. Solid lines delineate heterotrait-monomethod triangles. Dashed lines delineate heterotrait-heteromethod triangles. Values in both triangles are disciminant validity coefficients. E = aesthetics; S = safety; and C = capacity.

Table 5.3

Coherence and Performance Validity Matrices for Posner's and Anderson's Studies

Coherence validity matrix

		Method			
		Posner (perception)		Anderson (verbal)	
Method	Concept	Concrete	Abstract	Concrete	Abstract
Posner (perception)	Concrete	(?)			
	Abstract	P	(?)		
Anderson (verbal)	Concrete	?	?	(?)	
	Abstract	?	?	A	(?)

Performance validity matrix

Criteria				
Concrete[a]	P	P	A	A
Abstract[b]	?	?	?	?

Note. Parentheses denote reliabilities. Ellipses are convergent validity coefficients. Triangles show discriminant validity coefficients, solid lines indicating heteroconcept-monomethod correlations and dashed lines indicating heteroconcept-heteromethod correlations in the coherence validity matrix and in the performance validity matrix, correlations between judgments of a concept by a particular method and the criterion for a different concept. P = value proved by Posner's (1969) study. A = value provided by Anderson's (1974) study. ? = values not provided.

[a]Criterion for "concrete" is reaction time. [b]Criterion for "abstract" is accuracy (used by Kolers, 1979).

Conclusion

There is clearly a growing uneasiness and dissatisfaction with the conventional methodology used in psychological research. Those, like Moray, who wish to apply the results of research, are increasingly expressing frustration over the dubious value of the products of laboratory research for solving problems of the workplace. Others, like Lamiell, who wish to study the role of personality in behavior, are increasingly expressing frustration over the slow pace of change in research methods. Those interested in the study of the effect of stress on J/DM cannot ignore these methodological issues. A defense of one's methodology should now be made explicit. It should no longer be sufficient merely to announce, "An ANOVA was performed."

Would the idiographic-statistical method combined with the representative design of experiments advance our research capability and competence over the present nomothetic-systematic approach? A positive, detailed answer to the question is put forward in Part III of the present work.

6. A New Approach

"Stress" now means so many different things to so many different researchers that each one must now say what s/he intends it to mean. For example, Hobfoll (1989) criticizes current efforts by stating that his goal is to clarify "the nature of what has proved to be a heuristic but vague construct." He challenges "current conceptualizations of stress . . . as being too phenomenological and ambiguous, and consequently, not given to direct empirical testing." He further argues that "researchers have tended to avoid the problem of defining stress, choosing to study stress without reference to a clear framework." But Hobfoll then continues the trend he has just criticized by introducing "a new stress model called the model of conservation of resources . . . as an alternative" (p. 513). These efforts typify the never-ending pursuit of elusive fragments of a concept. Integration of all these bits and pieces is so highly improbable that it is time to remove this word from the scientific literature. Therefore, we urge that researchers abandon the term *stress:* It should be replaced by a theory of cognition that describes and predicts functional relations between special environmental conditions of interest and subsequent cognitive activity. Abandoning the term "stress" should not be considered unusual; scientific progress often leads to the abandonment of conventional terms of use in the lay person's language and the development or invention of new, more differentiated, more specific terms useful for scientific investigation. In this section we present an outline of a psychological theory that has no need for the term "stress."

Disruption of Constancy:
An Outline of a Research-Based Theory of Stress

The theory to be presented here rests on two broad premises: (a) All organisms attempt to maintain *stable* relations with their environment, and (b) it is the *disruption* of stabilized relations that produces the behaviors of interest to the present reader.

The *first* premise can be refined, made specific and operational, and noncircular measures can be developed by reference to the empirically well-established phenomenon of *constancy*, that is, the achievement of the correct inference of a covert, distal variable despite changing overt, proximal information about it. The *second* premise leads directly to testable propositions regarding cognitive responses to a disruption of an organism's stable relation (constancy) with its environment. And that is how our topic is to be defined. Thus, the ambiguities in the lay concepts of stress, and the numerous vague, and often circular, professional and scientific concepts of stress (noted by Hobfoll above) can be set aside in favor of concentration on a specific, well-known, critical form of behavior—constancy. The goal of this shift is to remove the confusing connotations of the word "stress" and to replace them with a focus on a specific form of behavior within the context of empirically well-established phenomena.

Constancy is one of the fundamental discoveries of psychology, if not *the* fundamental discovery. Constancy has been found to exist over a wide range of circumstances, including size, shape, color, brightness, density, weight, volume, tactile-kinesthetic sensations, loudness, and even values (see any text on perception). Brunswik

(1956) generalized the concept of constancy to mean "stabilization" of relations between organism and environment and observed that "stabilization effects of the kind studied in the thing constancies are of the very essence of life" (p. 23). More specifically, "stabilized relationships with the environment are biologically useful adjustments, especially when they anchor organismic orientation to properties of more or less remote, more or less vitally relevant solid objects of potential manipulation and locomotion such as landmarks, tools, enemies, or prey, which themselves are usually fairly stable or predictable" (p. 23). Therefore, it is surely plausible to argue that the *disruption* of constancy, that is, disruption of stabilization, presents a threat to the organism that requires not merely an affective response, but a cognitive one that is intended to re-establish stability and thus survival. (See Mandler, 1982; Simon, 1979, pp. 33-36 of *Models of Thought* for related ideas.)

Disruption of constancy therefore will be taken as a theoretical point of departure that offers freedom from the wide variety of atheoretical common sense suppositions regarding stress and judgment, as well as the numerous idiosyncratic definitions of stress that remain unique to specific investigators. Most important, because this approach embeds the situations and behaviors of interest in a well-known, well-documented, domain-independent aspect of human behavior in general, and human judgment in particular, investigators are thus afforded the opportunity to develop a cumulative rather than splintered discipline; researchers can build on empirical facts.

In addition, the concept "disruption of constancy" is researchable; it can readily be reduced to testable predictions regarding observable cognitive responses to manipulable, environmental, "independent" variables without the risk of circularity so frequently observed in the conventional definition of stress. And that, in turn, makes it possible to understand and predict judgment and decision making processes that are evoked when people are seeking to maintain constancy—a stable relation with their environment—but when circumstances prevent, hinder, or disrupt their efforts to do so.

Correspondence Constancy Versus Coherence Constancy

The above remarks attempt to persuade the reader that the introduction of the concept "disruption of constancy" represents a theoretical advance and an empirically useful point of departure. So far, however, this discussion has been limited to *correspondence constancy;* that is, constancy that is measured in terms of empirical accuracy, or the correspondence between judgments and the relevant properties of the person's environment. It is this form of constancy that so readily lends itself to discussion in terms of "stabilization," "survival," and the like; it is wholly *performance*-oriented.

As pointed out (see Section 2 above), however, there are two main topics within J/DM research: (a) performance (treated as correspondence constancy above) *and* (b) coherence. Therefore it is necessary to consider "coherence constancy." The latter refers to the ability to understand correctly the underlying logic or syntax of a problem despite its various forms of display or presentation. This

ability can be readily observed in those instances when the same problem is presented either in substantive or symbolic form.

Can the concept of "coherence constancy" be set in parallel to the concept of "correspondence constancy"? No one has ever *explicitly* claimed that coherence constancy is related to "stabilization," "survival," or has claimed that coherence constancy should have the same conceptual status as correspondence constancy. Nevertheless, there is reason to believe that, in postagrarian if not postindustrial society, coherence constancy, that is, getting the logic of the problem correct despite its various forms of presentation, is indeed survival related. The environment is constructed, engineered so as to induce, if not demand, analytical cognition far more than in the time of the hunter/gatherer. Calculation, for example, is far more frequently demanded today than even 500 years ago. Thus, coherence constancy is as valuable, probably, as correspondence constancy, at least in western civilization.

No one would dispute the research that shows high achievement of correspondence constancy empirically observed in homo sapiens. On the other hand, the results of much of the research in J/DM seem to show that coherence constancy is poor. Indeed it is the lack of coherence constancy that leads to the above-described gloomy conclusions regarding the quality of human judgment. But it is often claimed that poor coherence constancy doesn't matter, or doesn't matter much: Kahneman and Tversky, the authors of much of the research that casts doubt on human cognitive abilities related to coherence, suggest that, although biases (i.e., errors) are frequent,

heuristics are robust, that is, they provide answers not far from the logically correct answer, and, therefore, human beings are not threatened in their existence by poor coherence constancy—as they would be if they had poor correspondence constancy. Just how robust various heuristics are remains uncertain. Nevertheless, although coherence constancy is taken to be far worse than correspondence constancy, its consequences have so far been judged by researchers not to be as severe as a lack of correspondence constancy.

Disruption of Constancy and Subsequent Cognitive Behavior: Some Examples

Although the occurrence of constancy phenomena of both types has been widely studied, the *disruption* of constancy has rarely been investigated (aside from the "ganzfeld" and sensory deprivation experiments which do not cast much light on the cognitive processes affected).

Disruption of constancy will be met by a variety of cognitive responses; some, of course, will be appropriate and thus useful, others not. The nature of the demands made by the environment at the time of the disruption will also vary widely. For example, restoration of constancy may *demand creativity* or it may require the *suppression of creativity*, depending on whether the source of disruption is endogenous or exogenous.

Endogenous Factors and Demand for Creativity

The flight deck crew of UAL 232 faced complete loss of control of the airplane by customary means (loss of aileron, rudder, and

elevator control) which required them to control the airplane only by changes in engine speeds. Such an endogenous disruption had never been encountered before in the history of civil aviation; it demanded that a new method had to be found almost immediately; in short, *creativity was demanded immediately.* The change in the endogenous properties of the task had disrupted the ability to maintain control of vertical and horizontal orientation and thus threatened the existence of all concerned. Constancy, or orientation was reduced to a perceptual-motor task of directing the airplane by means of throttle manipulation; analytical cognitive efforts to solve the problems ended once the crew had (rapidly) determined what had happened.

Disruption of Constancy, Ambient Exogenous Factors, and Suppression of Creativity

In contrast with UAL 232, the crew of the Aloha airliner (which lost a significant part of its fuselage) was required to *suppress* creativity; that is, they were required to employ *normal* procedures for landing—despite terrifying exogenous factors (noise, decompression, etc.). The successful suppression of creativity—new, untried activity—prevented disaster (behavior also true of the crew of the United airliner that lost a cargo door shortly after leaving Honolulu). In short, disruption of constancy can require entirely different cognitive activity depending on task circumstances. The hypothesis offered here is that endogenous change requires *creativity on demand;* exogenous change requires the *suppression of creativity.*

Creativity on Demand

The study of judgment under disruption of constancy will therefore require further consideration of the process of creativity on demand (very rapid problem solving), as well as suppression of creativity. That is, when constancy is reduced sufficiently for the person to notice the loss (i.e., acquire outcome feedback), then cognitive activity will be required to re-establish constancy. In the case of endogenous change the person's (a) environmental supports (colleagues, handbooks, etc.) and (b) personal resources (knowledge, experience, training) may provide the necessary cognitive means to regain constancy rapidly, if not effortlessly. If so, then, of course, the loss of constancy requires only the motivation necessary to resort to the use of these resources to repair the disruption. If, however, neither set of resources enables one to remedy the loss of constancy, then creativity will be demanded. And such creativity may draw largely on intuitive cognition—a point we discuss below.

But if *exogenous* factors cause disruption, then environmental supports and personal resources will have to be relied upon to assist the operator to defy the disrupting agents and to continue to perform as if they did not exist. (See Lee & Bussolari, 1989, for a report on the response of pilots with limited experience to platform motion in a simulator.)

Fear of Error

The two examples just presented are those of experts (skilled pilots) functioning under extreme changes in endogenous and exogenous task properties that disrupted constancy. We turn now to a

third example of how changes in *social* constancy may induce fear of error of the consequences of making the wrong choice (creativity—or its suppression are not an issue in this example). When there is uncertainty, wrong choices will be of two kinds—false positives and false negatives. Choosing between these errors in highly uncertain conditions presents a difficult, serious problem for cognitive activity, as may be seen in the following example.

In November, 1991, a Russian officer was forced to choose between taking his men over to the originators of the November 1991 coup (a bad mistake if the coup failed) or to take them over to those fighting the coup (a bad mistake if the coup succeeded). The situation is described by Cullen in the *New Yorker* magazine as follows:

> On the following day, the Russian forces began suborning the rest of the troops in the city. Dmitri Rosnin joined an agitprop group that set out in a private car from the [Russian] White House [where Boris Yeltsin presided]. "We went looking for columns of armor," he told me. "Around the Rossiya Theatre we saw a unit of about eighty armored vehicles. We started to talk with the company commander—to find out what his assignment was, and to argue that it was an illegal power he was serving. The soldiers in the unit quite happily took and read the documents we were passing out—Yeltsin's statement, and so on. I didn't see any efforts to stop them, nor did I see any loaded weapons. I told the commander that I myself was a former company commander. I also told him a lie—that some storm troops had come to the White House and had come over to our side. He said he knew about that. I said that if I were in his position I would take my men and go over to Yeltsin's side. I noticed that he had tears in his eyes. I understood what kind of stress he was under. He couldn't reply. I said, 'O.K., if you can't do that, then at least explain to the soldiers the situation—tell them that Yazov is a criminal, and Pugo, too. You're a Russian officer, not just a Soviet officer. Have you forgotten that?' He didn't answer. He just listened, and his eyes got red. I said, 'O.K., then just don't shoot us.' As I turned to go, I heard his reply: 'O.K.' " (Cullen, 1991, pp. 76-77)

The Russian officer reporting this story clearly recognized that the officer in charge of the tank column was under "stress" because he faced a difficult cognitive problem: He was being required to estimate the dangers associated with a false positive (saying "yes" when he should be saying "no") on the one hand, and a false negative (the opposite) on the other hand, with terrible consequences following either error.

Choosing between dichotomous circumstances under uncertainty is always difficult; but because the tank commander had no reliable information of his own (in fact, the raconteur admits he gave him *false* information), the task was completely uncertain. The risk of error was surely great, enough one might say, to make a strong man weep.

Loss of Constancy and Predictions from Cognitive Continuum Theory

We now provide an example of how a theory of judgment can be employed to further our understanding and predictions of cognitive activity in response to disruption. Cognitive Continuum Theory makes explicit references to task conditions and differentiates them according to task properties specified a priori. Most important, CCT makes different predictions regarding the cognitive activity that is induced in relation to different task conditions. Therefore, application of CCT to the disruption of constancy should enable the investigator to differentiate among task conditions and, as a result, to offer predictions of cognitive activity in response thereto. (Details regarding CCT can be found in Hammond, 1988; Hammond et al., 1987.)

Loss of coherence constancy. It is generally assumed that experts work in task conditions designed and engineered to be compatible with coherence, and which, therefore encourage and induce analytical cognition; that is, Cell 9 is the normal situation for experts (see Table 6.1 below). (See Vicente, 1990.)

What happens when the unexpected occurs, when coherence is lost, and when there is no trained remedial response available (e.g., UAL 232)? If there is time pressure for solution, CCT predicts that cognition will be driven, will move, from Cell 9 toward Cell 1 *(intuition* will be employed).

Table 6.1

Task Continuum Index

Cognitive Continuum Index	I	9	A
I	1 **best**	2 mediocre	3 poor
9	4 mediocre	5 **best?**	6 mediocre
A	7 mediocre risk of large error	8 mediocre	9 **best**

Is impairment, or degradation of performance inevitable? No; performance may *improve* if the environmental task conditions that disrupted constancy change so that intuitive cognition corresponds more closely to the task conditions than analytical cognition would.

More specifically, if (a) the task conditions move to Cell 1 and (b) the subject is driven to Cell 1 by time pressure or the failure of analytical cognition and thus cannot employ an analytical model, then performance will *improve*. The robust properties of the weighted average linear model organizing principle will be employed in precisely the situations in which it provides high correspondence constancy. Specifically, if the task conditions change so as to provide greater redundancy and increased linearity, it will be easier for the operator to be right for the wrong reason.

We believe these circumstances could account for the improvement in performance of the NCAR weather forecasters (described above) under conditions of increased weather activity. As the storm activity increased, more and more redundant cues appeared, and, as time pressure increased, cognitive activity was driven toward intuition. These environmental task properties induced the use of a robust organizing principle (e..g, a weighted averaging procedure), thus producing more accurate forecasts.

Loss of correspondence constancy. Loss of correspondence constancy will have various cognitive consequences (rather than the predictable consequence from loss of coherence constancy). The properties of the task will not move in a predictable direction from Cell 1. The task may continue to be intuition inducing, may acquire some analysis-inducing properties or may become completely analysis-inducing. CCT predicts different cognitive consequences, depending on the nature of the change in task properties and the cognitive response. For example, if, as in a navigation situation, a

visual/perceptual environment fails, that is, no longer allows correspondence constancy, and if the navigator must now rely on displays that provide only a few cues, then analytical cognition will be induced. That is, cognition will be driven from Cell 1 to Cell 9, with attendant cognitive consequences (see Table 6.2).

Table 6.2

Task Continuum Index

Cognitive Continuum Index	I	Q	A
I	1 **best**	2 mediocre	3 poor
Q	4 mediocre	5 **best?**	6 mediocre
A	7 mediocre risk of large error	8 mediocre	9 **best**

In short, CCT emphasizes the need for careful analysis of the change in task properties that occurs when the unexpected occurs. The task properties at t_1 (prior to the time the unexpected occurs) *and* the task properties of the new situation at t_2 (the time the unexpected occurs) need to be examined. According to CCT, they will predict the cognitive activity in both situations and the behavioral consequences.

Conclusion

The purpose of presenting the outlines of a theory of stress based on constancy phenomena was to indicate how such a theory could be developed within the context of a well-established, well-documented, domain-independent aspect of human behavior in general and human judgment in particular. Our principal effort was to indicate that there is no need for special theories of stress; indeed, we urge the abandonment of that term from scientific discourse. It has served only to distract and has resulted in the accumulation of large but diverse, disorganized, noncumulative literature. (See Part I of this Series.)

Following the argument for a research endeavor based on the *disruption* of constancy, we provided examples of how a theory of judgment (CCT) would provide predictions of cognitive activity in response to the changed conditions that disrupt either correspondence constancy or coherence constancy.

Although the arument presented here is sketchy and perhaps vague, it pretends to provide no more than an indication of how a theory of stress can be rooted in well-documented constancy phenomena. Nevertheless, we believe that this approach could very well lead to a research program that could draw on the insights achieved by the current research programs in the field of J/DM.

7. General Conclusions

Here we briefly present conclusions drawn from each of the previous sections.

Current Topics

Two topics ([a] the competence of human judgment in the achievement of rationality, and [b] the performance of human judgment in the achievement of accuracy) have attracted numerous researchers from various research programs. Different, incommensurable methods and theories have been employed, and different conclusions about rationality and performance have been reached. There is no agreement among researchers in the field regarding the competence and performance of human judgment. This is a regrettable situation for stress researchers, inasmuch as it deprives them of an agreed-upon baseline from which to measure the effects of stress. No resolution of this problem appears to be imminent.

Current Textbooks and Articles

Textbooks. The minimal amount of research on the effects of stress on J/DM (noted in Part I) is made evident by the almost complete absence of significant discussion of the topic in the several textbooks and anthologies written by authorities in the J/DM field. The only discussion is given by Yates (1990), but that is only two pages long. Thus, the conclusion drawn in Part I is supported by the treatment of this topic by the authors of the J/DM textbooks; evidently they didn't find enough material to write about. It should be pointed

out, however, that the human factors researchers have produced numerous articles on the effects of stress—usually in relation to exogenous factors (e.g., heat, noise, etc.)—on the performance of human judgment, not its rationality, that are completely ignored by the authors of J/DM textbooks. The absence of references to this material in the textbooks indicates a gap between the two fields that should be closed. Human factors researchers, however, rarely discuss the cognitive processes involved in judgment and decision making that occupy the major part of the J/DM textbooks.

The articles reviewed here show a similar gap: Research articles in J/DM rarely consider stress, though a few examine the effects of mood; articles in the human factors field rarely discuss process.

Current Theories and Models of J/DM

Although researchers in the field of J/DM have paid little or no attention to the effects of stress on J/DM, they have developed theories and models of J/DM processes that may well be very useful in relation to this topic. The fact that there are persistent, stable research programs in this field, each of which is demonstrating cumulative results, suggests that these programs offer excellent potential for the study of stress and judgment. But if progress is to be achieved, some external effort will need to be made to induce the members of these programs to consider the topic of stress; there is no indication that any of the J/DM research programs are interested in this topic.

Methodology

The problem of generalization of results continues to pose difficulties. Researchers—particularly in the human factors field—are increasingly calling for greater representiveness of research situations, and a greater emphasis on naturalistic research. There is also a continuing insistence on the need for the study of individuals rather than the comparison of group averages. Considerable tension remains between those who advocate the methodology that persists in conventional experimental psychology and those who advocate representative research conditions and idiographic methodology. There is a growing difference in methodological standards between those whose work is dominated by conventional methods and those who favor abandoning them in favor of naturalistic, idiographic research. No reconciliation of these different methodological issues is apparent.

A New Approach

First, we urge the abandonment of the word stress by researchers; it is a term that has outlived its usefulness to them. Second, we argue for a research endeavor based on a well-known, well-documented phenomenon—constancy. We argue that the disruption of constancy provides a concrete, empirical basis for examining the interrelation of cognition and affect—which is the fundamental issue. Third, Cognitive Continuum Theory was briefly presented to show that a theory of judgment could be used to indicate circumstances in which performance will improve as well as become impaired. Examples were provided from military situations as well as

civil aviation. The development of this conceptual framework—and its methodological implications—might well provide the new approach so clearly needed. We make this attempt in Part III of this series.

References

Abelson, R. P., & Levi, A. (1985). Decision making and decision theory. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 1, pp. 231-309). New York: Random House.

Adelman, L., & Bresnick, T. (1991). *Examining the effect of information sequence on expert judgment: An experiment with Patriot air defense officers using the Patriot air defense simulator* (Tech. Rep.). Reston, VA: Decision Science Consortium, Inc.

Adelman, L., Tolcott, M. A., & Bresnick, T. A. (1990). *Examining the effect of information order on expert judgment* (Tech. Rep.). Reston, VA: Decision Sciences Consortium, Inc.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice-Hall.

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt.

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1991a). Is human cognition adaptive? *The Behavioral and Brain Sciences, 14,* 471-517.

Anderson, J. R. (1991b). Optimality and human memory. *The Behavioral and Brain Sciences, 14,* 215-216.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review, 96,* 703-719.

Anderson, N. H. (1962). Application of an additive model to impression formation. *Science, 138*, 817-818.

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology, 70*, 394-400.

Anderson, N. H. (1979). Algebraic rules in psychological measurement. *American Scientist, 67*, 555-563.

Anderson, N. H. (1981). *Foundations of Information Integration Theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of Information Integration Theory*. New York: Academic Press.

Arkes, H. R., & Hammond, K. R. (Eds.). (1986). *Judgment and decision making: An interdisciplinary reader*. Cambridge: Cambridge University Press.

Asare, S. K. (1990). *The auditor's going concern opinion decision: Interaction of task variables and the sequential processing of evidence* (Tech. Rep.). Gainesville, FL: University of Florida, Fisher School of Accounting.

Ashton, A. H., & Ashton, R. H. (1988). Sequential belief revision in auditing. *The Accounting Review, 63*, 623-641.

Ashton, R. H., & Ashton, A. H. (1990). Evidence-responsiveness in professional judgment: Effects of positive versus negative evidence and presentation mode. *Organizational Behavior and Human Decision Processes, 46*, 1-19.

Ballou, D. P., & Pazer, H. L. (1990). A framework for the analysis of error in conjunctive, multi-criteria, satisficing decision processes. *Decision Sciences, 21*, 752-770.

Barker, R. G. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior.* Stanford, CA: Stanford University Press.

Baron, J. (1988). *Thinking and deciding.* Cambridge: Cambridge University Press.

Bartlett, F. C. (1932). *Remembering.* London: Cambridge.

Bazerman, M. H. (1990). *Judgment in managerial decision making* (2nd ed.). New York: Wiley.

Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica, 47*, 89-104.

Brehmer, B. (1987). Models of diagnostic judgements. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), *New technology and human error.* London: Wiley.

Brehmer, B., Allard, R., & Lind, M. (1983). *Fire-fighting: A paradigm for the study of dynamic decision making.* Paper presented at the Ninth Research Conference on Subjective Probability, Utility, and Decision Making, Groningen, The Netherlands.

Brehmer, B., & Dörner, D. (1991). *Experiments with computer simulated microworlds: Escaping the narrow straits of the laboratory as well as the deep blue sea of the field study.* Uppsala, Sweden: Uppsala University, Department of Psychology.

Brehmer, B., & Joyce, C. R. B. (Eds.). (1988). *Human judgment: The SJT view.* Amsterdam: Elsevier.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review, 64*, 205-215.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon.

Broadbent, D. E. (1971). *Decision and stress.* London: Academic Press.

Bronfenbrenner, U. (1974). *Ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Bruner, J. S., Goodnow, J., & Austin, G. A. (1956). *A study of thinking.* New York: Wiley.

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review, 50*, 255-272.

Brunswik, E. (1952). The conceptual framework of psychology. In *International encyclopedia of unified science* (Vol. 1, no. 10, pp. 4-102). Chicago: University of Chicago Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193-217.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.

Burrell, Q. L., & Cane, V. R. (1982). The analysis of library data. *Journal of the Royal Statistical Society, Series A, 145*, 439-471.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Casscells, W., Schoenberger, A., & Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299*, 999-1000.

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Christensen-Szalanski, J. J. J., & Beach, L. R. (1984). The citation bias: Fad and fashion in the judgment and decision literature. *American Psychologist, 39*, 75-78.

Cosmides, L., & Tooby, J. (1990, August). *Is the mind a frequentist?* Paper presented at the 31st Annual Meeting of the Psychonomics Society, New Orleans, LA.

Cullen, R. (1991, November 4). Report from Moscow: The coup. *The New Yorker*, pp. 54, 56, 58-59, 62-82, 84-88.

Dawes, R. M. (1988). *Rational choice in an uncertain world.* San Diego: Harcourt, Brace, Jovanovich.

de Keyser, V., Decortis, F., & van Daele, A. (1988). The approach of Francophone ergonomy: Studying new technologies. In V. de Keyser, T. Qvale, B. Wilpert, & S. A. R. Quintanilla (Eds.), *The meaning of work and technological options.* Chichester: Wiley.

Doerner, D. (1987). On the difficulties people have in dealing with complexity. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), *New technology and human errors.* London: Wiley.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 183-201.

Ebbesen, E. B., & Konecni, V. J. (1980). On the external validity of decision-making research: What do we know about decisions in the real world? In T. Wallsten (Ed.), *Cognitive processes in choice and decision behavior*. Hillsdale, NJ: Erlbaum.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 41*, 380-417.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmutz (Ed.), *Formal representation of human judgment* (pp. 17-52). New York: Wiley.

Edwards, W. (in press). Discussion: Of human skills. *Organizational Behavior and Human Performance.*

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*, 395-416.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32*, 53-88.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin, 99*, 3-19.

Einhorn, H. J., & Hogarth, R. M. (1987). *Adaptation and inertia in belief updating* (Tech. Rep.). University of Chicago: Center for Decision Research.

Entin, E. E., & Serfaty, D. (1990). *Information gathering and decisionmaking under stress* (TR-454). Burlington, MA: Alphatech.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1-38). Cambridge: Cambridge University Press.

Evans, J. S. B. T. (1984). In defense of the citation bias in the judgment literature. *American Psychologist, 39*, 1500-1501.

Fechner, G. T. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). New York: Holt, Rinehart, & Winston. (Original work published 1860)

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50*, 123-129.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422-444). Cambridge: Cambridge University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Gigerenzer, G. (1991a). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98*, 254-267.

Gigerenzer, G. (1991b). How to make cognitive illusions disappear: Beyond "heuristics and biases." *European Review of Social Psychology, 2*, 83-115.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 513-525.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98*, 506-528.

Gilligan, S. G., & Bower, G. H. (1984). Cognitive consequences of emotional arousal. In C. E. Izard, J. Kagan, & R. B. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 547-588). Cambridge: Cambridge University Press.

Ginossar, Z., & Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and Social Psychology, 52*, 464-474.

Graesser, C. C., & Anderson, N. H. (1974). Cognitive algebra of the equation: Gift size = generosity x income. *Journal of Experimental Psychology, 103*, 692-699.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics (Reprint)*. Huntington, NY: Krieger. (Original work published 1966)

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics, 95*, 537-557.

Hacking, I. (1986). Culpable ignorance of interference effects. In D. MacLean (Ed.), *Values at risk* (pp. 136-154). Totowa, NJ: Rowman and Allanheld.

Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review, 62*, 255-262.

Hammond, K. R. (1966). *The psychology of Egon Brunswik*. New York: Holt, Rinehart, and Winston.

Hammond, K. R. (1978). *Psychology's scientific revolution: Is it in danger?* (Tech. Rep. No. 211). Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1980). *The integration of research in judgment and decision theory* (Tech. Rep. No. 226). Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1986a). Generalization in operational contexts: What does it mean? Can it be done? *IEEE Transactions on Systems, Man, and Cybernetics, SMC-16*, 428-433.

Hammond, K. R. (1986b). *A theoretically based review of theory and research in judgment and decision making* (Tech. Rep. No. 260). University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1988). Judgment and decision making in dynamic tasks. *Information and Decision Technologies, 14*, 3-14.

Hammond, K. R. (1990a). *The effects of stress on judgment and decision making: An overview and arguments for a new approach* (Tech. Rep. No. 320). Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1990b). Functionalism and illusionism: Can integration by usefully achieved? In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 227-261). Chicago, IL: University of Chicago Press.

Hammond, K. R., Hamm, R. M., & Grassia, J. (1986). Generalizing over conditions by combining the multitrait-multimethod matrix

and the representative design of experiments. *Psychological Bulletin, 100,* 257-269.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-17,* 753-770.

Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). *Human judgment and decision making: Theories, methods, and procedures.* New York: Hemisphere/Praeger.

Hammond, K. R., & Summers, D. A. (1972). Cognitive control. *Psychological Review, 79,* 58-67.

Hancock, P. A. (Ed.). (1987). *Human factors psychology.* Amsterdam: Elsevier.

Hancock, P. A., & Chignell, M. H. (1987). Adaptive control in human-machine systems. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 305-345). Amsterdam: Elsevier.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors, 31,* 519-537.

Hansen, R. D., & Donoghue, J. M. (1977). The power of consensus: Information derived from one's own and others' behavior. *Journal of Personality and Social Psychology, 35,* 294-302.

Harvey, L. O., Jr., Hammond, K. R., Lusk, C. M., & Mross, E. F. (in press). The application of signal detection theory to weather forecasting behavior. *Monthly Weather Review.*

Hastie, R. (1991). A review from a high place: The field of judgment and decision making as revealed in current textbooks. *Psychological Science, 2,* 135-138.

Hebb, D. O. (1949). *The organisation of behavior.* New York: Wiley.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44,* 513-524.

Hockey, [G.] R. [J.] (1979). Stress and the cognitive components of skilled performance. In V. Hamilton & D. M. Warburton (Eds.), *Human stress and cognition* (pp. 141-177). Chichester: Wiley.

Hockey, [G.] R. [J.] (1984). Varieties of attentional state: The effect of the environment. In R. S. Parasuraman & D. R. Davies (Eds.), *Varieties of attention.* Orlando, FL: Academic Press.

Hockey, G. R. J. (1986). Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and performance: Vol. 2. Cognitive processes and performance* (pp. 44-1—44-49). New York: Wiley.

Hockey, G. R. J., Gaillard, A. W. K., & Coles, M. G. H. (Eds.). (1986). *Energetics and human information processing.* Dordrecht: Nijhoff.

Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. *Psychological Bulletin, 90,* 197-217.

Hogarth, R. M. (1987). *Judgement and choice: The psychology of decision* (2nd ed.). Chichester: Wiley.

Hogarth, R. M., & Einhorn, H. J. (1989). *Order effects in belief updating: The belief-adjustment model* (Tech. Rep.). University of Chicago: Center for Decision Research.

Hogarth, R. M., McKenzie, C. R. M., Gibbs, B. J., & Marquis, M. A. (1991). Learning from feedback: Exactingness and incentives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 734-752.

Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review, 71*, 42-60.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment.* New York: Free Press.

Jensen, R. S., & Adrion, J. (1985). *Aeronautical decision-making for instrument pilots.* Columbus, OH: Aviation Research Associates.

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology, 45*, 20-31.

Jorna, P. G. A. M., & Visser, R. T. B. (1991). Selection by flight simulation: Effects of anxiety on performance. In E. Farmer (Ed.), *Human resource management in aviation: Proceedings of the XVIII WEAAP conference: Vol. 1.* Aldershot, England: Avebury Technical.

Jungermann, H. (1983). The two camps on rationality. In R. W. Schulz (Ed.), *Decision-making under uncertainty* (pp. 63-86). Amsterdam: Elsevier.

Kahneman, D. (1991). Judgment and decision making: A personal view. *Psychological Science, 2,* 142-145.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263-291.

Klein, G., Orasanu, J., & Calderwood, R. (Eds.). (in press). *Decision making in action: Models and methods.* Norwood, NJ: Ablex.

Kleinmuntz, D. N. (1987). Human decision processes: Heuristics and task structure. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 123-157). Amsterdam: Elsevier.

Kolers, P. A. (1979). A pattern analyzing basis of recognition. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 363-384). Hillsdale, NJ: Erlbaum.

Lakatos, I. (1981). History of science and its rational reconstructions. In I. Hacking (Ed.), *Scientific revolutions* (pp. 107-127). Oxford: Oxford University Press.

Lamiell, J. T. (1982). The case for an idiothetic psychology of personality: A conceptual and empirical foundation. *Progress in Experimental Personality Research, 11,* 1-64.

Lamiell, J. T., & Trierweiler, S. J. (1986). Interactive measurement, idiothetic inquiry, and the challenge to conventional "nomotheticism." *Journal of Personality, 54*, 460-469.

Lang, P. J. (1979). Language, image and emotion. In P. Pliner, K. R. Plankstein, & I. M. Speigel (Eds.), *Perception of emotion in self and others (Vol. 5)*. New York: Plenum.

Larrick, R. P., Morgan, J. N., & Nisbett, R. E. (1990). Teaching the use of cost-benefit reasoning in everyday life. *Psychological Science, 1*, 362-370.

Laudan, L. (1981). A problem-solving approach to scientific progress. In I. Hacking (Ed.), *Scientific revolutions* (pp. 144-155). Oxford: Oxford University Press.

Lazarus, R. S. (1966). *Psychological stress and the coping process.* New York: McGraw-Hill.

Lee, A. T., & Bussolari, S. R. (1989). Flight simulator platform motion and air transport pilot training. *Aviation, Space, and Environmental Medicine, 60*, 136-140.

Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist, 43*, 431-442.

Lehman, D. R., & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology, 26*, 952-960.

Leon, M., & Anderson, N. H. (1974). A ratio rule from integration theory applied to inference judgments. *Journal of Experimental Psychology, 102*, 27-36.

Lewin, K. (1935). *A dynamic theory of personality.* New York: McGraw-Hill.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 551-578.

Lopes, L. L. (1976). Model-based decision and inference in stud poker. *Journal of Experimental Psychology: General, 105,* 217-239.

Lorenz, K. (1966). *On aggression.* New York: Harcourt, Brace, and World.

Lusk, C. M. (in press). Assessing components of judgments in an operational setting: The effects of time pressure on aviation weather forecasting. In O. Svenson & J. Maule (Eds.), *Time pressure and stress in human judgement and decision making.* New York: Plenum.

Lusk, C. M., & Hammond, K. R. (1991). Judgment in a dynamic task: Microburst forecasting. *Journal of Behavioral Decision Making, 4,* 55-73.

Lusk, C. M., Stewart, T. R., Hammond, K. R., & Potts, R. J. (1990). Judgment and decision making in dynamic tasks: The case of forecasting the microburst. *Weather and Forecasting, 5,* 627-639.

Mandler, G. (1975). *Mind and emotion.* New York: Wiley.

Mandler, G. (1982). Stress and thought processes. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (pp. 88-104). New York: Free Press.

Mann, L., & Janis, I. (1982). Conflict theory of decision making and the expectancy-value approach. In N. T. Feather (Ed.), *Expectations and actions: Expectancy-value models in psychology* (pp. 341-364). Hillsdale, NJ: Erlbaum.

Mano, H. (1990a). Anticipated deadline penalties: Effects on goal levels and task performance. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 154-172). Chicago: University of Chicago Press.

Mano, H. (1990b). *Judgments under distress: Assessing the role of unpleasantness and arousal in judgment formation* St. Louis, MO: Washington University, John M. Olin School of Business.

Massaro, D. W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review, 97*, 225-252.

Meehl, P. E. (1954). *Clinical vs. statistical prediction.* Minneapolis, MN: University of Minneapolis Press.

Meehl, P. (1990). *Corroboration and verisimilitude: Against Lakatos's "sheer leap of faith."* Minneapolis, MN: University of Minnesota, Minnesota Center for Philosophy of Science.

Meisser, W. F., Jr. (1990). *The sequencing of audit evidence: Its impact on the extent of audit testing and report formulation* (Tech. Rep.). Gainesville, FL: University of Florida, Fisher School of Accounting.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Mithen, S. J. (1990). *Thoughtful foragers: A study of prehistoric decision making.* Cambridge: Cambridge University Press.

Moray, N. (1969). *Attention: Selective process in vision and hearing.* London: Hutchinson.

Moray, N. (1990). Designing for transportation safety in the light of perception, attention, and mental models. *Ergonomics, 33,* 1201-1213.

Neisser, U. (1978). Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3-24). London: Academic Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nickerson, C. A. E., & McClelland, G. H. (1991). *A note on within-person versus across-persons tests of social-psychological models of the attitude-behavior relationship.* Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Nisbett, R. E., & Borgida, E. (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology, 32,* 932-943.

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science, 238,* 625-631.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339-363.

Orasanu, J., & Connolly, T. (in press). The reinvention of decision making. In G. Klein, J. Orasanu, & R. Calderwood (Eds.), *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Payne, J. W. (1976). Task complexity and contingent decision processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*, 366-387.

Payne, J. W. (1980). Information processing theory: Some concepts and methods applied to decision research. In T. S. Wallsten (Ed.), *Cognitive processes in choice and decision behavior*. Hillsdale, NJ: Erlbaum.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin, 92*, 382-402.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534-552.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43*, 87-131.

Payne, J. W., & Braunstein, M. L. (1978). Risky choice: An examination of information acquisition behavior. *Memory & Cognition, 5*, 554-561.

Phillips, L. D. (1983). A theoretical perspective on heuristics and biases in probabilistic thinking. In P. C. Humphreys, O. Svenson, & A. Vari (Eds.), *Analysing and aiding decision processes* (pp. 525-543). Amsterdam: Elsevier.

Piaget, J. (1981). *Intelligence and affectivity: Their relationship during child development.* Palo Alto, CA: Annual Reviews.

Raby, M., Wickens, C. D., & Marsh, R. (1990). *Investigation of factors comprising a model of pilot decision making: Part I. Cognitive biases in workload management strategy* (ARL-90-7/SCEEE-90-1). Urbana-Champaign, IL: University of Illinois, Aviation Research Laboratory.

Rachlin, H. (1989). *Judgment, decision, and choice.* New York: W. H. Freeman.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty.* Reading, MA: Addison-Wesley.

Raiffa, H. (1969). *Preferences for multiattributed alternatives* (Report No. RM-5868-DOT/RC). Santa Monica, CA: Rand Corporation.

Rasmussen, J. (1986). *Information processing and human machine interaction: An approach to cognitive engineering.* Amsterdam: Elsevier.

Reason, J. (1990). *Human error.* Cambridge: Cambridge University Press.

Rimm, D. C., & Somervill, J. W. (1977). *Abnormal psychology.* New York: Academic Press.

Rohrmann, B., Beach, L. R., Vlek, C., & Watson, S. R. (Eds.). (1989). *Advances in decision research: Edited proceedings of the 11th Conference on Subjective Probability, Utility, and Decision Making, Cambridge, UK, August 23-27, 1987.* Amsterdam: Elsevier.

Rothstein, H. G. (1986). The effects of time pressure on judgment in multiple cue probability learning. *Organizational Behavior and Human Decision Processes, 37*, 83-92.

Russo, J. D., & Schoemaker, P. J. H. (1989). *Decision traps: Ten barriers to brilliant decision-making and how to overcome them.* New York: Doubleday.

Sackett, D. L. (1989). Inference and decision at the bedside. *Journal of Clinical Epidemiology, 4*, 309-316.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 65*, 379-399.

Schoemaker, P. J. H. (1991). The quest for optimality: A positive heuristic of science? *The Behavioral and Brain Sciences, 14*, 205-245.

Schwartz, D. R., & Howell, W. C. (1985). Optional stopping performance under graphic and numeric CRT formatting. *Human Factors, 27*, 33-444.

Serfaty, D., Entin, E., & Tenney, R. (1989). Planning with uncertain and conflicting information. In S. E. Johnson & A. H. Levis (Eds.), *Science of command and control: Coping with complexity* (pp. 91-100). Fairfax, VA: AFCEA International Press.

Shanteau, J. (1974). Component processes in risky decision making. *Journal of Experimental Psychology, 103*, 680-691.

Simmel, E. C., Cerkovnik, M., & McCarthy, J. E. (1989). Sources of stress affecting pilot judgment. *Aviation, Space, and Environmental Medicine, 60,* 53-55.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99-118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63,* 129-138.

Simon, H. A. (1979). *Models of thought.* New Haven, CT: Yale University Press.

Simon, H. A. (1983). *Reason in Human Affairs.* Stanford, CA: Stanford University Press.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology, 28,* 1-39.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance, 6,* 649-744.

Smith, J. F., & Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin, 109,* 472-489.

Spielberger, C. (1979). *Understanding stress and anxiety.* New York: Harper and Row.

Stewart, T. R., Middleton, P., Downton, M., & Ely, D. (1984). Judgments of photographs vs. field observations in studies of perception and judgment of the visual environment. *Journal of Environmental Psychology, 4,* 283-302.

Stewart, T. R., Middleton, P., & Ely, D. W. (1983). Urban visual air quality judgments: Reliability and validity. *Journal of Environmental Psychology, 3,* 129-145.

Stewart, T. R., Moninger, W. R., Grassia, J., Brady, R. H., & Merrem, F. H. (1989). Analysis of expert judgment in a hail forecasting experiment. *Weather and Forecasting, 4,* 24-34.

Stokes, A. F. (1991). MIDIS—A microcomputer flight decision simulator. In E. Farmer (Ed.), *Human resource management in aviation: Proceedings of the XVIII WEAAP conference: Vol. 1* (pp. 107-121). Aldershot, England: Avebury Technical.

Stokes, A., Belger, A., & Zhang, K. (1990). *Investigation of factors comprising a model of pilot decision making: Part II. Anxiety and cognitive strategies in expert and novice aviators* (ARL-90-8/SCEEE-90-2). Urbana-Champaign, IL: University of Illinois, Aviation Research Laboratory.

Stokes, A. F., & Raby, M. (1989). Stress and cognitive performance in trainee pilots. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (Vol. 2, pp. 883-887). Santa Monica, CA: The Human Factors Society.

Stritter, E. (1977). *File migration.* Unpublished Ph.D. thesis (STAN-CS-77-594), Stanford University.

Svenson, O., Edland, A., & Slovic, P. (1990). Choices and judgments of incompletely described decision alternatives under time pressure. *Acta Psychologica, 75,* 153-169.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science, 182,* 990-1000.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin, 110,* 67-85.

Telfer, R. (1989). Pilot decision making and judgement. In R. S. Jensen (Ed.), *Aviation psychology* (pp. 154-175). Aldershot: Gower Technical.

Thurstone, L. L. (1931). The measurement of change in social attitudes. *Journal of Social Psychology, 2,* 230-235.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281-299.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124-1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293-315.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review, 93,* 371-384.

Vicente, K. J. (1990). Coherence- and correspondence-driven work domains: Implications for systems design. *Behaviour & Information Technology, 9,* 493-502.

Vicente, K. J., & Rasmussen, J. (in press). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics.*

Vlek, C. (1984). What constitutes "a good decision"? In K. Brocherding (Ed.), *Research perspectives on decision-making under uncertainty.* Amsterdam: Elsevier.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research.* Cambridge: Cambridge University Press.

Watson, S. R., & Buede, D. M. (1987). *Decision synthesis: The principles and practice of decision analysis.* Cambridge: Cambridge University Press.

Wells, G. L., & Harvey, J. H. (1977). Do people use consensus information in making causal attributions? *Journal of Personality and Social Psychology, 35,* 279-293.

Wickens, C. D. (1984). *Engineering psychology and human performance.* Columbus, OH: Merrill.

Wickens, C. D. (1987). Attention. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 29-80). Amsterdam: Elsevier.

Wickens, C. D. (in press). *Engineering psychology and human performance* (2nd ed.). New York: Harper Collins.

Wickens, C., & Flach, J. (1988). Human information processing. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 111-155). San Diego: Academic Press.

Wickens, C. D., Stokes, A., Barnett, B., & Hyman, F. (1988). *The effects of stress on pilot judgement in a MIDIS simulator* (Tech. Rep. No. ARL-88-5/INEL-88-1). Urbana-Champaign: University of Illinois, Aviation Research Laboratory, Institute of Aviation.

Williams, A. C., & Hopkins, C. O. (1958). *Aspects of pilot decision making.* Columbus, OH: OSU Aviation Psychology Laboratory.

Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology, 59*, 555-561.

Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice-Hall.

Yule, G. U. (1921). Critical notice [Review of *The essentials of mental measurement*]. *British Journal of Psychology, 12*, 100-107.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151-175.

Zajonc, R. B., & Markus, H. (1984). Affect and cognition: The hard interface. In C. E. Izard, J. Kagan, & R. B. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 73-101). Cambridge: Cambridge University Press.

Appendix

Berliner, H. J., Goetsch, G., Campbell, M. S., & Ebeling, C. (1990). Measuring the performance potential of chess programs. *Artificial Intelligence, 43*, 7-20.

Birnbaum, M. H., & Sotoodeh, Y. (1991). Measurement of stress: Scaling the magnitudes of life changes. *Psychological Science, 2*, 236-243.

Broadbent, D. E. (1980). The minimization of models. In A. J. Chapman & D. M. Jones (Eds.), *Models of man* (pp. 113-128). Leicester, England: The British Psychological Society.

Broadbent, D. E. (1990). Effective decisions and their verbal justification. In D. E. Broadbent, J. Reason, & A. Baddeley (Eds.), *Human factors in hazardous situations: Proceedings of a Royal Society Discussion Meeting held on 28 and 29 June 1989* (pp. 45-54). Oxford: Clarendon Press.

Camerer, C. F., & Johnson, E. J. (1991). The process-performance paradox in expert judgment: How can experts know so much and predict so badly? In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 195-217). Cambridge: Cambridge University Press.

Cohen, L. J. (1991). Intuition in surgery: Has it a legitimate role? *Theoretical Surgery, 6*, 98-99.

Coren, S., & Aks, D. J. (1991). Prediction of task-related arousal under conditions of environmental distraction. *Journal of Applied Social Psychology, 21*, 189-197.

Corrie, S. J. (1989). *Cockpit voice recorder group chairman's factual report of investigation* (DCA-89-M-A063). Washington, DC: National Transportation Safety Board, Bureau of Technology.

Dalgleish, L. I., & Drew, E. C. (1989). The relationship of child abuse indicators to the assessment of perceived risk and to the court's decision to separate. *Child Abuse and Neglect, 13*, 491-506.

Doyle, J. K. (1991). *Judgments of cumulative risk.* Unpublished doctoral dissertation, University of Colorado, Boulder.

Eddy, D. M. (1989). The confidence profile method: A Bayesian method for asessing health technologies. *Operations Research, 37*, 210-228.

Edland, A. (1989). *On cognitive processes under time stress: A selection review of the literature on time stress and related stress* (Supplement 68). Stockholm, Sweden: The University of Stockholm, Department of Psychology.

Elstein, A. S. (1991). On the psychology of clinical intuition. *Theoretical Surgery, 6*, 95-97.

Fox, R. (1991). Intuition: Pattern recognition. *Theoretical Surgery, 6*, 107.

Gleicher, F., Kost, K. A., Baker, S. M., Strathman, A. J., Richman, S. A., & Sherman, S. J. (1990). The role of counterfactual thinking in judgments of affect. *Personality and Social Psychology Bulletin, 16*, 284-295.

Green, R. (1990). Human error on the flight deck. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 327*, 503-512.

Gross, R., & Lorenz, W. (1990). Intuition in surgery as a strategy of medical decision making: Its potency and limitations. *Theoretical Surgery, 5*, 54-59.

Hamm, R. M. (1988). Moment-by-moment variation in experts' analytic and intuitive cognitive activity. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-18*, 757-776.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1986). Direct manipulation interfaces. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New perspectives on human-computer interaction* (pp. 87-124). Hillsdale, NJ: Erlbaum.

Itoh, J.-I., Yoshimura, S., Ohtsuka, T., & Masuda, F. (1990). Cognitive task analysis of nuclear power plant operators for man-machine interface design. In *Proceedings of the Topical Meeting on Advances in Human Factors Research on Man/Computer Interactions: Nuclear and Beyond* (pp. 96-102). La Grange Park, IL: American Nuclear Society.

Jensen, R. S. (Ed.). (1989). *Proceedings of the Fifth International Symposium on Aviation Psychology. Vol. 1.* Columbus, OH: The Ohio State University, The Aviation Psychology Laboratory.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*, 374-419.

Kastner, M., Harss, C., & Beerman, L. (1989). Measuring stress of helicopter pilots: An analysis of deficiencies in critical flight situations. In R. S. Jensen (Ed.), *Proceedings of the Fifth International Symposium on Aviation Psychology* (Vol. 2, pp. 539-

544). Columbus, OH: The Ohio State University, The Aviation Psychology Laboratory.

Keren, G. B., & Raaijmakers, G. W. (1988). On between-subjects versus within-subjects comparisons in testing utility theory. *Organizational Behavior and Human Decision Processes, 41*, 233-247.

Kieras, D. (in press). Diagrammatic displays for engineered systems: Effects on human performance in interacting with malfunctioning systems. *International Journal of Man-Machine Studies.*

Kleinmuntz, D. (1991). Decision making for professional decision makers. *Psychological Science, 2*, 135, 138-141.

Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106*, 395-409.

Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist, 36*, 276-289.

Lee, A. T. (1991). Aircrew decision-making behavior in hazardous weather avoidance. *Aviation, Space, and Environmental Medicine, 62*, 158-161.

Lynn, S. G., & Kellermann, A. L. (1991). Critical decision making: Managing the emergency department in an overcrowded hospital. *Annals of Emergency Medicine, 20*, 287-292.

Maher, J. W. (1991, May). *Why pilots are least likely to get good decision making precisely when they need it most.* Paper presentated at the meeting of the Sixth International Aviation Psychology Symposium, Ohio State University.

McPeek, B. (1991). Intuition as a strategy of medical decision making. *Theoretical Surgery, 6*, 83-84.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1990). Similarity involving attributes and relations: Judgments of similarity and difference are not inverses. *Psychological Science, 1*, 64-69.

Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry, 1*, 108-141.

Moray, N. (1987). Intelligent aids, mental models, and the theory of machines. *International Journal of Man-Machine Studies, 27*, 619-629.

Moray, N. (1988). *A lattice theory of mental models of complex systems* (EPRL-88-08). Urbana, IL: University of Illinois at Urbana-Champaign, Engineering Psychology Research Laboratory.

Moray, N. (1991). *Technosophy and humane factors: The future of ergonomics?* Urbana, IL: University of Illinois at Urbana-Champaign, Department of Mechanical and Industrial Engineering and of Psychology.

Morecroft, J. D. W. (1983). System dynamics: Portraying bounded rationality. *OMEGA: The International Journal of Management Science, 11*, 131-142.

Mosier-O'Neill, K. L. (1989). A contextual analysis of pilot decision making. In R. S. Jensen (Ed.), *Proceedings of the Fifth International Symposium on Aviation Psychology* (Vol. 1, pp. 371-376). Columbus, OH: The Ohio State University, The Aviation Psychology Laboratory.

National Transportation Safety Board. (1990). *Aircraft accident reports: Brief format, U.S. civil and foreign aviation, issue number 13 of 1988 accidents* (NTSB/AAB-89-20). Washington, DC: National Transportation Safety Board.

National Transportation Safety Board. (1991a). *Aircraft accident report: Northwest Airlines, Inc. Flights 1482 and 299 runway incursion and collision, Detroit Metropolitan/Wayne County Airport, Romulus, Michigan, December 3, 1990* (NTSB/AAR-91/05). Washington, DC: National Transportation Safety Board.

National Transportation Safety Board. (1991b). *Aircraft accident report: Runway collision of Eastern Airlines Boeing 727, Flight 111 and Epps Air Service Beechcraft King Air A100, Atlanta Hartsfield International Airport, Atlanta, Georgia, January 18, 1990* (NTSB/AAR-91/03). Washington, DC: National Transportation Safety Board.

Paunonen, S. V., & Jackson, D. N. (1986). Nomothetic and idiothetic measurement in personality. *Journal of Personality, 54*, 447-459.

Rasmussen, J. (1991). *Diagnostic reasoning in action.* Roskilde, Denmark: Risø National Laboratory.

Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, & The PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 45-76). Cambridge, MA: MIT Press.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & The PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 318-362). Cambridge, MA: MIT Press.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models* (pp. 7-57). Cambridge, MA: MIT Press.

Smith, C. A. P. (1990). *Decision making under time pressure: The effects of time pressure on information search strategy, decision strategy, consistency, and outcome quality.* Unpublished doctoral dissertation, The University of Arizona.

Te'eni, D. (1990). Direct manipulation as a source of cognitive feedback: A human—computer experiment with a judgement task. *International Journal of Man—Machine Studies, 33,* 453-466.